

AD_____

Award Number: DAMD17-02-1-0424

TITLE: Molecular Mechanisms of Breast Cancer Metastasis

PRINCIPAL INVESTIGATOR: Lori L. Wallrath, Ph.D.

CONTRACTING ORGANIZATION: University of Iowa
Iowa City, IA 52242

REPORT DATE: May 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20051101 084

REPORT DOCUMENTATION PAGEForm Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY
(Leave blank)**2. REPORT DATE**
May 2005**3. REPORT TYPE AND DATES COVERED**
Final (25 Apr 2002 - 24 Apr 2005)**4. TITLE AND SUBTITLE**

Molecular Mechanisms of Breast Cancer Metastasis

5. FUNDING NUMBERS

DAMD17-02-1-0424

6. AUTHOR(S)

Lori L. Wallrath, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)University of Iowa
Iowa City, IA 52242**E-Mail:** Lori-wallrath@iowa.edu**8. PERFORMING ORGANIZATION
REPORT NUMBER****9. SPONSORING / MONITORING
AGENCY NAME(S) AND ADDRESS(ES)**U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012**10. SPONSORING / MONITORING
AGENCY REPORT NUMBER****11. SUPPLEMENTARY NOTES****12a. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE**13. ABSTRACT (Maximum 200 Words)**

The five-year survival rate for women with a primary breast tumor is 98%. In contrast, the five-year survival rate of individuals with breast cancer metastasis is 26%. Understanding the molecular mechanisms of metastasis are essential for the prediction and prevention of metastatic disease. Studies of the chromosome associated protein termed Heterochromatin Protein 1-alpha (HP1^{Hsa}) will provide insights on the molecular changes that accompany breast cancer metastasis. HP1^{Hsa} is down-regulated in highly invasive/metastatic breast cancer cells compared with poorly invasive/non-metastatic breast cancer cells. This down-regulation is also observed in metastatic patient tissues. We have demonstrated that increasing the levels of HP1^{Hsa} in highly invasive/metastatic breast cancer cells causes a reduction in their invasive ability. Furthermore, reducing levels of HP1^{Hsa} in poorly invasive/non-metastatic breast cancer cells increase their invasion ability. Our working model is that HP1^{Hsa} alters the expression of genes required for invasion/metastasis. We are currently identifying the genes regulated by HP1^{Hsa}. Products of such genes will offer new molecules for the diagnosis and treatment of metastatic disease.

14. SUBJECT TERMS
metastasis**15. NUMBER OF PAGES**
64**16. PRICE CODE****17. SECURITY CLASSIFICATION
OF REPORT**
Unclassified**18. SECURITY CLASSIFICATION
OF THIS PAGE**
Unclassified**19. SECURITY CLASSIFICATION
OF ABSTRACT**
Unclassified**20. LIMITATION OF ABSTRACT**
Unlimited

Table of Contents

Cover.....	1
SF 298.....	2
Table of Contents.....	3
Introduction.....	4
Body.....	4
Key Research Accomplishments.....	8
Reportable Outcomes.....	9
Conclusions.....	9
References.....	10
Appendices.....	11

INTRODUCTION

Breast cancer accounts for approximately 30% of all diagnosed cancer and is the most common malignancy in women. The lethality of breast cancer is mainly attributed to its ability to metastasize throughout the body. According to the American Society, the five-year survival rate for individuals with metastatic breast cancer is 26%, in contrast to 98% for those with non-metastatic breast cancer. Although there has been significant progress made in the identification and understanding of breast cancer over recent years, there remains a need for the identification of molecular markers that definitively distinguish poorly invasive/non-metastatic tumors from highly invasive/metastatic tumors. To identify such prognostic/predictive markers, a better understanding of the molecular events leading to breast cancer progression is needed.

We have identified a molecular marker that is a candidate breast cancer metastasis suppressor. This marker, termed Heterochromatin Protein 1 (HP1^{Hsα}), is an evolutionarily conserved non-histone chromosomal protein (9). HP1 primarily localizes to centric heterochromatin where it plays a role in chromosome segregation and silencing of genes brought into juxtaposition with heterochromatin (9). HP1 also localizes within the gene-rich euchromatic regions of the genome where it plays a role in gene regulation (2). We discovered that HP1^{Hsα}, one of three HP1 proteins in humans, was significantly down-regulated in highly invasive/metastatic breast cancer cells compared with poorly invasive/non-metastatic breast cancer cells (6). This observation was specific for HP1^{Hsα}, and not the other two HP1 family members, HP1^{Hsβ} and HP1^{Hsγ}. We also discovered a similar down-regulation of HP1^{Hsα} levels in breast cancer patients: HP1^{Hsα} was abundant in the nuclei of cells from primary breast tumors, but dramatically reduced in cells of metastatic tissues (6). Given the role of HP1 proteins in gene regulation, we hypothesize HP1^{Hsα} alters the expression of genes involved in invasion/metastasis.

BODY

Progress made within the past year on each task is described below. All tasks are identical to those stated in the approved STATEMENT OF WORK in the original proposal.

Task 1: Determine the consequences of HP1^{Hsα} expression on tumor metastasis markers and global transcriptional expression in human breast cancer cells.

This task is based on the observation that poorly invasive/non-metastatic breast cancer cells (MCF-7) have levels of HP1^{Hsα} similar to most cell types, whereas, highly invasive/metastatic cells (MDA-MB-231) have low levels HP1^{Hsα} (6). Given the role of HP1^{Hsα} in gene regulation (9), we hypothesize that alterations in the levels of HP1^{Hsα} result in changes in gene expression. We have experimentally modulated the levels of HP1^{Hsα} in both MCF-7 and MDA-MB-231 cells and assayed for changes in gene expression. An adenovirus delivery system was utilized to increase levels of HP1^{Hsα} in MDA-MB-231 cells. This viral delivery system was used due to our inability to isolate MDA-MB-231 cells that stably expressed an *EGFP-HP1^{Hsα}* transgene. In the last grant

period we demonstrated that expression of EGFP-HP1^{Hsc} in MDA-MB-231 reduced *in vitro* invasion by 30% relative to control cells expressing a nuclear tagged EGFP (NLS-GFP) (Figure 1). Expression of a mutant form of HP1^{Hsc} containing an amino acid substitution in the HP1^{Hsc} chromo shadow domain (I165E) that disrupts HP1 dimerization did not alter the invasive phenotype. [See Task 3 below for more information of HP1^{Hsc} domain structure.] In contrast, expression of a mutant form of HP1^{Hsc} carrying an amino acid substitution in the HP1^{Hsc} chromo shadow domain (W174A) that disrupts interactions with known protein partners containing a PxVxL motif reduced invasion to levels similar to those obtained following expression of EGFP-HP1^{Hsc}. Taken these data together, we conclude that increased levels of HP1^{Hsc} and dimerization of HP1^{Hsc} cause reduced invasion potential in MDA-MB-231 cells.

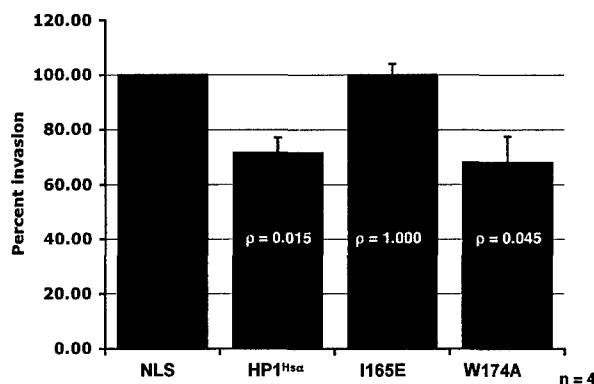


Figure 1. Expression of EGFP-HP1^{Hsc} reduces invasion of MDA-MB-231 cells. Cells were infected with adenovirus expressing either wild type or mutant forms of EGFP-HP1^{Hsc}. Results were compared to cells expressing an EGFP-NLS control. Invasion assays were performed according to published procedures (3).

In this grant cycle we tested for effects of HP1^{Hsc} on expression of candidate HP1-regulated genes. Candidate genes included those involved in cancer cell invasion and metastasis as well as known HP1-target genes. RNA from MDA-MB-231 cells infected with adenovirus expressing EGFP-HP1^{Hsc}, the I165E mutant form of HP1^{Hsc} or the W174A mutant form of HP1^{Hsc} was used for gene expression studies. The results were compared with those obtained from RNA isolated from cells expressing the NLS-GFP control and uninfected cells. Gene expression was measured by semi-quantitative RT-PCR and real-time PCR. Eight genes known to be associated with HP1 did not change in expression upon introduction of wild type or mutant forms of HP1^{Hsc} relative to control cells. Even though these genes have been shown to be associated with HP1^{Hsc}, the lack of change in MDA-MB-231 cells could be due to cell type specificity for HP1^{Hsc} association and/or compensation by other HP1 family members. Five tumor suppressor genes, including *BRCA1*, were investigated for changes in gene expression, but none of these genes changed in MDA-MB-231 cells upon introduction of HP1^{Hsc} relative to control cells. Twenty metastasis-associated genes, including ten integrin genes involved in cell-cell and cell-extracellular matrix interactions, and four matrix metalloproteases, were tested for alterations in gene expression. One metastasis-associated gene, *integrin α3* (*ITGA3*) showed altered expression upon changes in HP1^{Hsc} dosage (Figure 2). Expression of HP1^{Hsc} and the W174A mutant form of HP1^{Hsc} resulted a 3 to 6-fold up-regulation of *ITGA3* expression compared to uninfected MDA-MB-231 control cells.

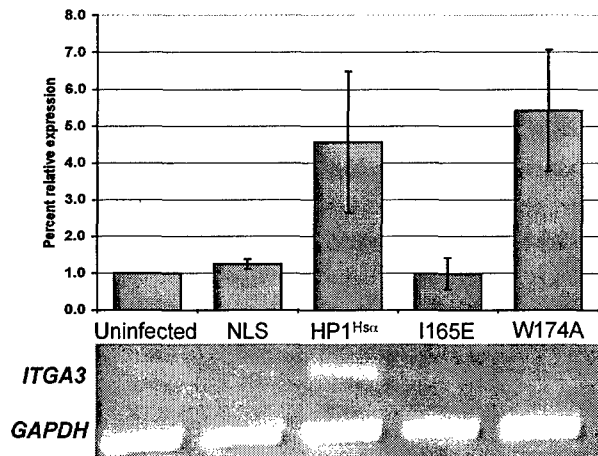


Figure 2. ITGA3 is up-regulated in response to increased levels of HP1^{Hsa} in MDA-MB-231 cells. Top represents a histogram of ITGA expression under each condition. The results from three independent samples are shown. Below are the semi-quantitative RT-PCR results for one sample.

Expression of the nuclear-tagged EGFP and the I165E mutant form of HP1 did not affect *ITGA3* expression compared to control cells. Therefore, *ITGA3* is up-regulated upon increased levels of HP1^{Hsc} in MDA-MB-231 cells. *ITGA3* expression has been correlated with tumor and metastasis suppression (ref), consistent with our model of HP1 as a metastasis suppressor.

As a second approach to modulating HP1^{Hsc} levels in breast cancer cells, we knock-down levels of HP1^{Hsc} using two different RNAi methods. First, HP1^{Hsc} levels were knocked-down using an adenovirus expressing an shRNAs against HP1^{Hsc}. We use adenovirus delivery when large numbers of cells are required for down-stream analysis, such as *in vitro* invasion assays (3). An adenovirus expressing shRNA to *GFP* was used as a non-specific control to rule out possible effects due to viral infection. Second, HP1^{Hsc} levels were knocked-down using small siRNAs purchased from Dharmacon. While the purchased siRNAs are expensive and not economical for large-scale experiments, the use of synthetic dsRNAs eliminates viral effects. Hence, the siRNAs were used for the global gene expression studies.

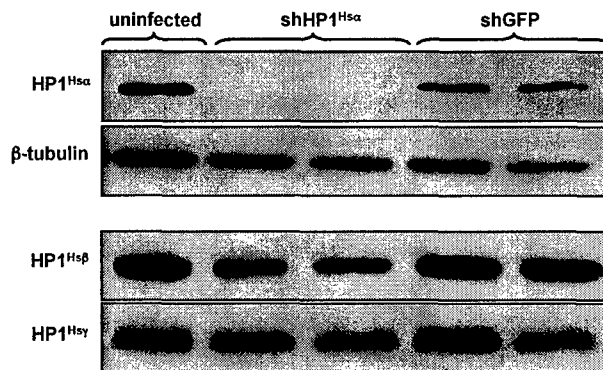


Figure 3. Knock-down of HP1^{Hsc} using shRNAs expressed from an adenovirus construct. An adenovirus expressing a shRNA against GFP was used as a negative control. Knock-down was specific for HP1^{Hsc}. Western analysis was performed 72 hours post infection.

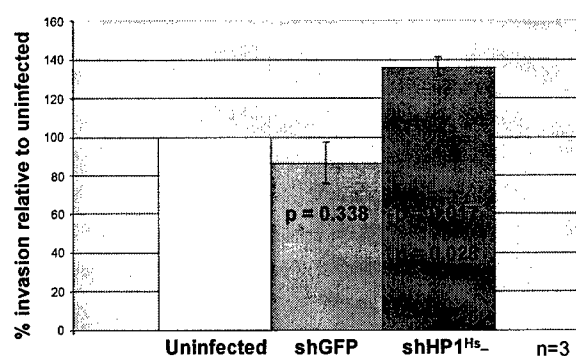


Figure 4. Knock-down of causes increased invasion of MCF-7 cells. Cells infected with an adenovirus expressed shGFP serve as controls.

Using shRNAs we obtained greater than 95% knock-down of HP1 in MCF-7 cells as evidenced by western analysis (Figure 3). Knock-down was specific for HP1^{Hsα}; levels of HP1^{Hsβ} and HP1^{Hsγ} did not dramatically change in the HP1^{Hsα} knock-down cells relative to controls (Figure 3).

To identify changes in gene expression that might account for the reduction in *in vitro* invasion upon expression of HP1^{Hsα}, we performed microarray analysis of RNA isolated from HP1^{Hsα} knock-down cells and cells transfected with a control scrambled dsRNA (Ambion). Microarray analysis has been performed on three independent experimental and control RNA samples and the results are currently being analyzed using Microsoft Suite and R. Approximately 200 genes show a 2-fold increase in expression upon knock-down of HP1^{Hsα}, and approximately 200 genes show a decrease in expression. These numbers are consistent with the data we obtained from gene expression studies of HP1 mutants in *Drosophila* (2). A list of the top 10 genes that increase in the HP1^{Hsα} knock-down cells relative to control cells is shown below (Table 1). We are in the process of performing statistical calculations on the microarray data, perform gene ontology analysis on the genes that change expression, and confirming changes in gene expression by RT-PCR. The data will provide insights on the molecular pathways that lead to invasion/metastasis.

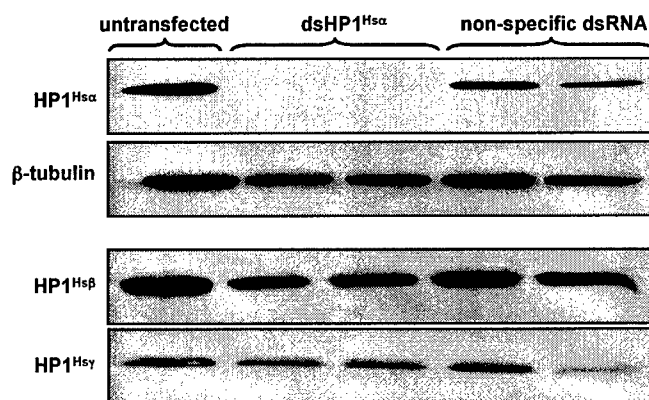


Figure 5. Knock-down of using dsRNA. Knock-down is specific for HP1^{Hsα}; levels of HP1^{Hsβ} and HP1^{Hsγ} are not dramatically altered. Western analysis was performed 72 hours after transfection with dsRNA.

Task 2: Determine the molecular mechanisms of $HP1^{Hsa}$ down-regulation in human breast cancer invasion/metastasis.

This task was completed and published during the previous year.

Task 3: Determine the domains of $HP1^{Hsa}$ required for invasion and metastasis.

$HP1^{Hsa}$ has a two-domain structure consisting of an amino chromo domain (CD) and a carboxy chromo shadow domain (CSD). The CD associates with methylated lysine 9 of histone H3 (5) and is thought to be the primary mechanism of HP1 localization within centric heterochromatin. The CSD homodimerizes (1); this dimerization generates a surface that interacts with a variety of proteins possessing a PxVxL pentapeptide motif (8, 10). In order to determine the domains of $HP1^{Hsa}$ involved in metastasis, we generated two mutant forms of $HP1^{Hsa}$. The first is an amino acid substitution in the α -helical region of the CD that disrupts the dimerization of HP1 (7). The second is an amino acid substitution within the CSD (W174A) that disrupts interactions with the pentapeptide motif-containing proteins (10). During the previous grant period, we demonstrated that expression of wild type $HP1^{Hsa}$ correlated with a 30% reduction in *in vitro* invasion relative to controls (Figure 1). A 30% reduction was also observed for the W174A mutation, suggesting that interactions with penta-peptide motif containing proteins do not play a role in regulating invasion. In contrast, there was no change in invasion following expression of the I165E, which inhibits dimerization relative to the control cells (Figure 1).

In this grant cycle we have addressed the effects of these mutants on candidate gene expression. *ITGA3* expression is altered upon introduction of wild type $HP1^{Hsa}$ and the W174A mutant, but not the I165E mutant (Figure 2). This data is consistent with the changes observed for *in vitro* invasion, where dimerization appears to be critical for decreasing invasion potential (Figure 1 and text of Task 1). *ITGA3* gene expression has been correlated with tumor and metastasis suppression (4). We are currently verifying the changes in *ITGA3* by real-time PCR and performing immunostaining to examine *ITGA3* protein levels.

KEY RESEARCH ACCOMPLISHMENTS

Obtained greater than 95% knock-down of $HP1^{Hsa}$ using adenoviral vectors expressing shRNAs directed against $HP1^{Hsa}$.

Demonstrated that MCF-7 cells expressing the shRNA against $HP1^{Hsa}$ have a 40% increased invasion potential than control shRNA expressing cells.

Obtained greater than 95% knock-down of HP1 in MCF-7 cells using short synthetic siRNAs.

Demonstrated that knock-down of $HP1^{Hsa}$ in MCF-7 cells does not dramatically alter the levels of $HP1^{Hsb}$ and $HP1^{Hsc}$.

Demonstrated that alterations in HP1^{Hsc} levels within MDA-MB-231 cells and MCF-7 cells does not alter their growth rate (data not shown), supporting HP1^{Hsc} as a metastasis suppressor, rather than a tumor suppressor.
 Demonstrated that increased expression of HP1^{Hsc} in MDA-MB-231 cells alters expression of *INTGA3*.
 Performed microarray analysis of MCF-7 with reduced levels of HP1^{Hsc}.

REPORTABLE OUTCOMES

Publication

1. Norwood, L.E., Grade, S.K., Cryderman, D.E., Hines, K.A., Furiasse, N., Toro, R., Li, Y., Dhasarathy, A., Kladde, M.P., Hendrix, M.J.C. and Kirschmann, D.A. and L.L. Wallrath (2004). Conserved properties of HP1^{Hsc}. *Gene* 335: 37-48.
2. Norwood, L.E., Hines, K.A., and L.L. Wallrath (2004) Deciphering the code of silence: Mechanisms of gene repression with connections to human disease. *ChemTracts* 17: 308-324.

Poster Abstracts

1. Moss, T.J., Sloat, S.L., Hendrix, M.J.C., Kirschmann, D.A. and L.L. Wallrath (2004). Effects of knocking-down Heterochromatin Protein 1 in breast cancer cells. University of Iowa Genetics Program Retreat, Iowa City, IA.
2. Moss, T.J., Sloat, S.L., Hendrix, M.J.C., Kirschmann, D.A. and L.L. Wallrath (2004).
3. The role of Heterochromatin Protein 1 in breast cancer cells. University of Iowa Medical Scientist Training Program Retreat, Backbone State Park, IA.
4. Norwood, L.E., Wright, L., Margaryan, N., Hendrix, M.J.C., Kirschmann, D.A. and L.L. Wallrath (2004) Heterochromatin Protein 1: Development of a novel breast cancer metastasis marker. Genetics PH.D. Interdisciplinary Program Retreat, Iowa City, IA.
5. Norwood, L.E., Moss, T.J., Margaryan, N., Sloat, S., Wright, L., Hendrix, M.J.C., Kirschmann, D.A. and L.L. Wallrath (2005). Heterochromatin Protein 1 (HP1^{Hsc}) as a candidate breast cancer metastasis suppressor. AACR 96th Annual Meeting, Anaheim, CA.
6. Moss, T.J., Sloat, S.L., Hendrix, M.J.C., Kirschmann, D.A. and L.L. Wallrath (2004). The effects of knock-down on breast cancer invasion/metastasis. University of Iowa, College of Medicine Research Week, Iowa City, IA.
7. Moss, T.J., Norwood, L.E., Margaryan, N., Cook, S.L., Wright, L., Hendrix, M.J.C., Dirschmann, D.A. and L.L. Wallrath (2005). Heterochromatin Protein 1 as a candidate breast cancer metastasis suppressor. DOD Era of Hope, Philadelphia, PA.

CONCLUSIONS

During the past year we have completed all of the tasks described in the original proposal. Our findings indicate that altered levels of HP1^{Hsc} in breast cancer cells leads to changes in invasion and gene expression. We have requested a "no cost extension" to cover the expenses associated with the following: (1) verification of the microarray

results by a second assay such as RT-PCR, (2) publication of a second manuscript on the new findings reported above (targeted for publication in Cancer Research) and (3) travel and lodging for the 2005 Era of Hope meeting (see Abstracts below). Our findings are consistent with the idea that HP1^{Hsc} is one of a small class of genes known as metastasis suppressor. Our research will be of great interest to investigators that work in the areas of breast cancer metastasis, chromatin structure and gene expression.

REFERENCES

1. Brasher, S. V., B. O. Smith, R. H. Fogh, D. Nietlispach, A. Thiru, P. R. Nielsen, R. W. Broadhurst, L. J. Ball, N. V. Murzina, and E. D. Laue. 2000. The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *Embo J* 19:1587-97.
2. Cryderman, D. E., S. K. Grade, Y. Li, L. Fanti, S. Pimpinelli, and L. L. Wallrath. 2005. Role of Drosophila HP1 in euchromatic gene expression. *Dev Dyn* 232:767-74.
3. Gehlsen, K. R., H. N. Wagner, Jr., and M. J. Hendrix. 1984. Membrane invasion culture system (MICS). *Med Instrum* 18:268-71.
4. Guo, W., and F. G. Giancotti. 2004. Integrin signalling during tumour progression. *Nat Rev Mol Cell Biol* 5:816-26.
5. Jacobs, S. A., and S. Khorasanizadeh. 2002. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295:2080-3.
6. Kirschmann, D. A., R. A. Lininger, L. M. Gardner, E. A. Seftor, V. A. Otero, A. M. Ainsztein, W. C. Earnshaw, L. L. Wallrath, and M. J. Hendrix. 2000. Down-regulation of HP1Hsalph expression is associated with the metastatic phenotype in breast cancer. *Cancer Res* 60:3359-63.
7. Lechner, M. S., G. E. Begg, D. W. Speicher, and F. J. Rauscher, 3rd. 2000. Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Mol Cell Biol* 20:6449-65.
8. Lechner, M. S., D. C. Schultz, D. Negorev, G. G. Maul, and F. J. Rauscher, 3rd. 2005. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain. *Biochem Biophys Res Commun* 331:929-37.
9. Li, Y., D. A. Kirschmann, and L. L. Wallrath. 2002. Does heterochromatin protein 1 always follow code? *Proc Natl Acad Sci U S A* 99 Suppl 4:16462-9.
10. Thiru, A., D. Nietlispach, H. R. Mott, M. Okuwaki, D. Lyon, P. R. Nielsen, M. Hirshberg, A. Verreault, N. V. Murzina, and E. D. Laue. 2004. Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin. *EMBO J* 23:489-99.

APPENDICES

1. Two Publications
2. Six Poster Meeting Abstracts

Conserved properties of HP1^{Hsα}

Laura E. Norwood^a, Stephanie K. Grade^a, Diane E. Cryderman^a, Karrie A. Hines^a,
Nicholas Furiasso^a, Rafael Toro^a, Yuhong Li^a, Archana Dhasarathy^b, Michael P. Kladde^b,
Mary J.C. Hendrix^c, Dawn A. Kirschmann^c, Lori L. Wallrath^{a,*}

^aThe Department of Biochemistry, The University of Iowa, 3136 MERF, Iowa City, IA 52242, USA

^bThe Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX, USA

^cThe Department of Anatomy and Cell Biology, The University of Iowa, Iowa City, IA, USA

Received 3 December 2003; received in revised form 3 March 2004; accepted 5 April 2004

Available online 5 June 2004

Received by A.J. van Wijnen

Abstract

Heterochromatin protein 1 Hsα (HP1^{Hsα}) is one of three human proteins that share sequence similarity with *Drosophila* HP1. HP1 proteins are enriched at centric heterochromatin and play a role in chromatin packaging and gene regulation. In humans, HP1^{Hsα} is down-regulated in highly invasive/metastatic breast cancer cells, compared to poorly invasive/non-metastatic breast cancer cells. To gain insight into this differential regulation, we have cloned the HP1^{Hsα} gene and characterized its genomic region. HP1^{Hsα} is located on human chromosome 12q13.13, 589 bp upstream of the divergently transcribed *hnRNP1* gene. Analysis of the promoter region revealed that differential regulation of HP1^{Hsα} between the two types of breast cancer cells is lost upon mutation of an USF/c-myc transcription factor binding site located 172 bp upstream of the predicted HP1^{Hsα} transcription start site. These findings provide insights into the down-regulation of HP1^{Hsα} in highly invasive/metastatic breast cancer cells. To examine the functional properties of HP1^{Hsα}, experiments were performed using *Drosophila melanogaster* as a genetic system. When human HP1^{Hsα} was expressed in transgenic *Drosophila*, silencing of reporter genes inserted at centric and telomeric locations was enhanced. Furthermore, expression of HP1^{Hsα} rescued the lethality of homozygous *Su(var)2-5* mutants lacking HP1. Taken together, these results demonstrate the participation of HP1^{Hsα} in silent chromatin formation and that HP1^{Hsα} is a functional homologue of *Drosophila* HP1.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Breast cancer metastasis; *Drosophila*; Gene silencing; Heterochromatin

1. Introduction

Heterochromatin protein 1 (HP1) was first discovered in *Drosophila melanogaster* and has since been found in a variety of eukaryotes from *Schizosaccharomyces pombe* to humans (Eissenberg and Elgin, 2000). *Drosophila*, mice and humans have three HP1 family members. All HP1 proteins

have a conserved amino domain termed the chromo domain (CD) and a carboxy domain termed the chromo shadow domain (CSD), separated by a less conserved hinge region (Eissenberg and Elgin, 2000). The HP1 CD binds to methylated lysine nine of histone H3; this interaction is important for HP1 localization at centric regions of chromosomes (Brehm et al., 2004).

The HP1 CSD homodimerizes, forming a site of interaction for several nuclear proteins possessing a pentapeptide motif (PxVxL), including the HP1 CSD itself (Cowieson et al., 2000; Smothers and Henikoff, 2000). The hinge region of some HP1 family members interacts with RNA and histone H1 (Nielsen et al., 2001; Muchardt et al., 2002). Thus, HP1 can be thought of as a bridging molecule that links various nuclear proteins to the chromosome.

Abbreviations: bp, base pair(s); kb, kilobase(s); HP1, heterochromatin protein 1; CD, chromo domain; CSD, chromo shadow domain; CAF1, chromatin assembly factor 1; µg, microgram; EGFP, enhanced green fluorescent protein; PCR, polymerase chain reaction; 5' RACE, 5' rapid amplification of cDNA ends; CMV, cytomegalovirus; USF, upstream stimulatory factor.

* Corresponding author. Tel.: +1-319-335-7920; fax: +1-319-384-4770.

E-mail address: lori-wallrath@uiowa.edu (L.L. Wallrath).

The known functions of HP1 proteins are largely based on genetic data. HP1 was identified in *Drosophila* mutagenesis screens for modifiers of heterochromatic gene silencing (Weiler and Wakimoto, 1995). Mutations in the *Drosophila* gene encoding HP1, *Su(var)2-5*, are homozygous lethal; heterozygotes show suppression of silencing of genes placed near heterochromatin, implying a role for HP1 in chromatin packaging. In *S. pombe*, mutations in the HP1-like protein Swi6 lead to chromosome segregation defects (Ekwall et al., 1995). Information gleaned from studies of HP1 in model organisms allows one to infer the function of HP1 homologs in mammalian cells, where genetic assays are not currently available.

In humans, *HP1^{Hsx}* is specifically down-regulated in highly invasive/metastatic breast cancer cells compared with poorly invasive/non-metastatic breast cancer cells, both at mRNA and protein levels (Kirschmann et al., 2000). Consistent with these cell culture phenotypes, staining of breast cancer tissue samples with antibodies to HP1^{Hsx} showed that HP1^{Hsx} was decreased in distant metastases compared to primary tumor tissues (Kirschmann et al., 2000). In this study, we identify sequences within the *HP1^{Hsx}* promoter region that are responsible for differential expression in metastatic vs. nonmetastatic breast cancer cell lines. In addition, we show here that HP1^{Hsx} is a functional homolog of *Drosophila* HP1.

2. Materials and methods

2.1. Isolation of genomic clones

A lambda genomic library made from the whole placenta of a 27-year-old healthy Caucasian female (Clontech) was screened to recover *HP1^{Hsx}* genomic clones. 1.5×10^6 independent clones were screened using standard techniques. *HP1^{Hsx}* cDNA was random prime labeled with ³²P-dATP (Amersham Multiprime labeling kit) and used for hybridization. Southern analysis was used to identify fragments within the clones containing *HP1^{Hsx}*. These fragments were isolated and sequenced to determine intron/exon boundaries of the *HP1^{Hsx}* gene.

2.2. 5' RACE

RNA was isolated from MCF-7 cells using TRIzol reagent (Life Technologies). RNA was amplified according to the 5' RACE System (Life Technologies) using a primer specific to a region 34–59 bp downstream of the stop codon. cDNA was amplified at 50°C instead of the typical 42°C to minimize secondary structure. The cDNA was PCR amplified using the abridged anchor primer (Life Technologies) specific to the C-tailed cDNA and a primer specific to the fifth exon. An additional extension cycle at 72°C for 3 min was added at the end of the PCR cycles. The PCR products were cloned (TA cloning system, Invitrogen) and sequenced.

2.3. Cells and culture conditions

MCF-7 cells were kindly supplied by Dr. F. Miller (Karmanos Cancer Institute, Detroit, MI). MDA-MB-231 cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA). Cell lines were maintained as previously described (Kirschmann et al., 1999).

2.4. Bisulfite genomic sequencing

Genomic DNA was isolated and analyzed by bisulfite genomic sequencing as previously described (Kladde et al., 1996). PCR products, amplified from bisulfite-deaminated DNA using JumpStart *Taq* DNA polymerase (Sigma), were purified and subjected to primer extension as described (Kladde et al., 1996), except that the final concentrations of dNTPs (A, C, T) and ddGTP were 50 and 150 μM, respectively. Exclusion of dGTP from the PCR product primer extension reactions generated high termination efficiencies (>96%) (Kladde et al., 1996) at template cytidines (nucleotides that were ^{5me}C residues in vivo).

2.5. Plasmid constructs

Fragments of the *HP1^{Hsx}* promoter region (positions –600, –466, –418, –286, –166, and –110 bp relative to transcription start at +1) fused to exon one (+143) were cloned into the pGL3-Basic vector (Promega). Forward primers used for generating the deletion constructs were as follows: –600 bp primer 5'-GCAGAAAGGAGC-GAGCTCACGAACGTATC-3', –466 bp primer 5'-CCTGCTATTGAGCTCTGGTGCCACATTGC-3', –418 bp primer 5'-GGTCGTTCTACGAGCTCTCCACC-3', –286 bp primer 5'-CTTCCACGAGCTCATATTACAGT-CAAG-3', –166 bp primer 5'-GTAAAATGGCGA-GCTCTGCGCA-3', –110 bp primer 5'-CGTGAAATG-GAGCTCAGGAGTAGG-3'. The reverse primer used for all of the *HP1^{Hsx}* promoter deletion constructs recognizes *HP1^{Hsx}* exon one and the pGL3-Basic vector: 5'-AGATCTCGAGCCCGGGATTGAGAGTGATCA-3'. The mutant transcription factor binding sites were generated using the QuikChange site-directed mutagenesis kit (Stratagene) and the following primers: δEF1 (mδEF1): 5'-CGCATTAAGAAGTTCCCCCCCCCTCTGAGAA-CACG-3' and 5'-CGTGTTCTCAGAGGGGGGGG-GAACTTCTTAATGCG-3', c-myb (mmyb): 5'-CGTTTTGGCGGGCCCCCCCCCTTGCGCAGAAGG-3' and 5'-CCTTCTGCGCAAGGGGGGGGGCCCCGCC-AAAACG-3', USF/c-myc site at position –109 (mUSFp): 5'-CCTCTGAGACCCCCCAATGGCGGGCAGGAG-TAG-3' and 5'-CTACTCCTGCCCGCCATTGGGG-GGGGTCTCAGAGG-3', USF/c-myc site at position –172 (mUSFd): 5'-CTCTTGTGACCGGGGG-GGGAGTAAAATGGCG-3' and 5'-CGCCATTT-TACTCCCCCCCCGGTCAACAAGAG-3'. The double mutants of mmyb and mUSFp (mmybUSFp) and mUSFp

and mUSFd (mUSFpUSFd) were also made using the above primers.

2.6. Transient transfection assays

MCF-7 and MDA-MB-231 cells were grown to 80% confluency. A total of 1 µg of DNA, including 0.5 µg of promoter construct and 0.5 µg of *CMV-lacZ* (kind gift of Dr. Andrew Russo), was transfected into the cell lines using Effectene Transfection Reagent (Qiagen). The cells were grown for 48 h, collected with Cell Culture Lysis Reagent (Promega), and assayed for luciferase and β-galactosidase expression. Luciferase expression was measured using Luciferase Assay Substrate (Promega) to monitor expression from *HP1^{Hsx}* promoter constructs. β-Galactosidase expression was measured using Galacto-Light Plus System (Applied Biosystems) to normalize for transfection efficiency. Light units were measured on a 96-well plate luminometer (Dynex). Normalized luciferase light unit measurements were set relative to light unit measurements obtained for a promoter construct containing 4 kb of *HP1^{Hsx}* upstream sequences, including exon one. This 4-kb construct gives uniform low levels of expression in both cell lines. These data were analyzed using the Microsoft Excel two samples unequal variance Student's *t*-test.

2.7. P-element construct and *Drosophila* germ line transformation

HP1^{Hsx} was fused in frame with *EGFP* and inserted into the P-element vector pCaSpeRhs-act (<http://thummel.genetics.utah.edu/>) containing an *hsp70* promoter to drive expression of the fusion gene and a mini-*white⁺* gene for selection of transformants. To generate an untagged *HP1^{Hsx}* construct, *HP1^{Hsx}* cDNA was inserted into the P-element vector pCaSpeR-hs-act. Both resulting P-element constructs were independently injected into *y, w^{67c23}* *Drosophila* embryos, along with P-turbo helper plasmid encoding transposase, according to standard germ-line transformation procedures. Daily heat-shock treatments lead to an estimated three-fold higher expression of *HP1^{Hsx}*-EGFP than the endogenous HP1 protein as judged by western analysis (data not shown).

2.8. *Drosophila* genetics

All *Drosophila* stocks were raised on standard corn meal sucrose media (Shaffer et al., 1994) at 25°C. Females with the genotype *P[w⁺, hsp70-HP1^{Hsx}]; Su(var)2-5⁰⁴/CyO, GFP* were crossed to males of the genotype *Su(var)2-5⁰²/CyO, GFP*. Crosses were heat shocked at 37°C for 45 min daily. Rescue of lethality was indicated by the presence of straight winged adults, representing the genotype *P[w⁺, hsp70-HP1^{Hsx}]; Su(var)2-5⁰⁴/Su(var)2-5⁰²*.

2.9. Immunostaining of polytene chromosomes

Third instar larvae were heat shocked at 37°C for 1 h and allowed to recover at room temperature for 2 h. Salivary glands were dissected, fixed, squashed and stained with a monoclonal antibody to HP1 (C1A9) and a polyclonal antibody to GFP (Molecular Probes) according to published procedures (Platero et al., 1995).

2.10. Northern analysis

RNA for Northern analysis was isolated from third instar larvae after heat shock at 37°C for 1 h according to published procedures (Wallrath et al., 1990). Levels of mRNA produced by the heterochromatic transgenes were measured by hybridization with *barley* cDNA sequences fused to the *hsp26* transgene and labeled with ³²P-dATP (Amersham) using random prime labeling (Amersham). An *rp49* cDNA was used as a control for RNA loading.

3. Results

3.1. Structure of the *HP1^{Hsx}* genomic region

HP1^{Hsx} is down-regulated in highly invasive/metastatic breast cancer cells in comparison to poorly invasive/non-metastatic breast cancer cells (Kirschmann et al., 2000). To better understand the mechanism of *HP1^{Hsx}* down-regulation, we have determined the structure of the *HP1^{Hsx}* genomic region, including the promoter region (Fig. 1). Clone F2-10 contains exons two and three surrounded by repetitive sequences typically found in introns. This clone spans a region approximately 9 kb upstream of exon two to 300 bp downstream of exon three. Exon two contains the methionine translation start codon. Clone F2-11 contains exons three, four, and five. A second screen, using sequences corresponding to exon one and 150 bp upstream as a probe, identified four clones containing the *HP1^{Hsx}* promoter region. Clone 3-4 contains 11 kb of the *HP1^{Hsx}* promoter region in addition to exon one that is 5' untranslated sequence. Taken together, *HP1^{Hsx}* is encoded by five exons spanning 38 kb.

5' RACE was performed using a primer to *HP1^{Hsx}* (positions +799 to +824) to identify the potential transcription start site. Three products, having their 5' ends within 22 bp of each other, were identified. We designated +1 as the 5' end of the longest 5' RACE product, extending exon one of *HP1^{Hsx}* an additional 54 bp upstream as compared to the NCBI CBX5 cDNA sequence NM_012117 (Fig. 1).

A bioinformatics analysis was performed on *HP1^{Hsx}* promoter region sequences. Using MatInspector V2.2 (<http://transfac.gbf.de/TRANSFAC/>) at stringent conditions (core sim 1.0, matrix sim 0.95), sequences from –601 to +143 were analyzed for known transcription factor binding

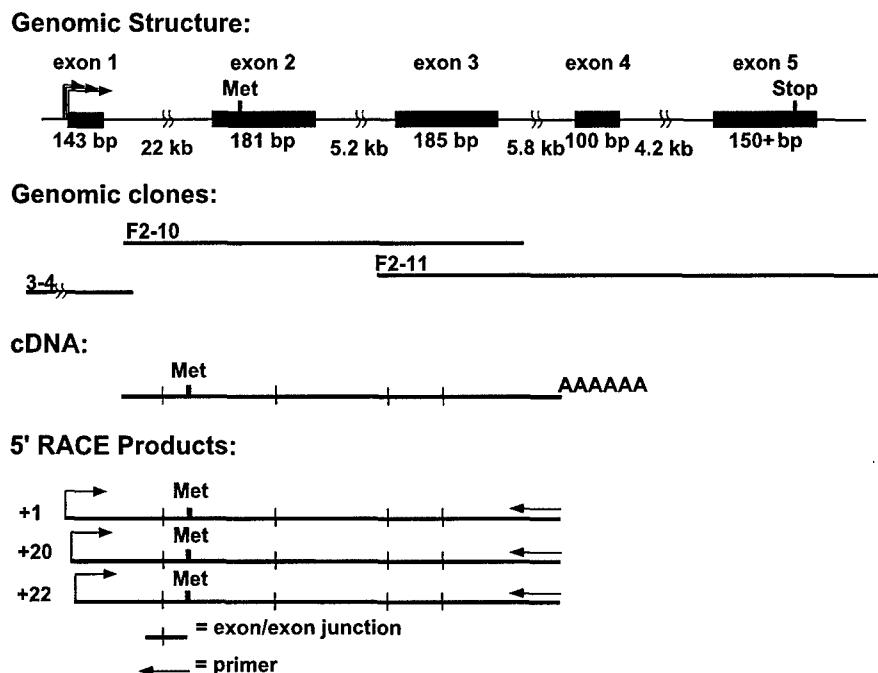


Fig. 1. The genomic structure of the human *HP1^{Hsx}* gene with the lengths of each exon and intron in bp or kb are shown. Transcription start sites are assigned on the basis of 5' RACE products. The translation methionine start (ATG) and stop are shown. The genomic clones are represented by lines under the corresponding genomic regions. Three clones obtained using 5' RACE start within 54 bp upstream of the cDNA sequences previously published in Genbank. The clone containing the most 5' sequence was designated +1; additional clones start at +20 and +22.

sites. Using these criteria, 37 binding sites, some overlapping with each other, were identified within this region, but no TATA box was identified. The lack of a TATA box is consistent with having multiple transcription start sites (Gum et al., 2003).

The *hnRNPA1* gene is divergently transcribed, starting at position –589 bp upstream of the *HP1^{Hsx}* transcription start (Fig. 3A). Although the promoter regions of *hnRNPA1* and *HP1^{Hsx}* are in close proximity, they appear to be independently regulated in breast cancer cells. Unlike the different levels of *HP1^{Hsx}* observed in the two breast cancer cell lines, *hnRNPA1* levels are unchanged between MDA-MB-231 and MCF-7 cells (data not shown). DNase I footprinting of the region between the *hnRNPA1* and *HP1^{Hsx}* genes was performed using HeLa cells (Biamonti et al., 1993). Six potential transcription factor binding sites found in our bioinformatics analysis of the *HP1^{Hsx}* promoter region correspond to the DNase I footprints previously identified, including two SP1 sites, two CAAT boxes, a CREB/c-jun site, two USF/c-myc sites, and a c-myc site (Biamonti et al., 1993) (Fig. 3A). A δ EF1 site within the promoter region is also of interest. The human AREB6 repressor protein that binds to the δ EF1 site is up-regulated in highly invasive/metastatic cell lines compared to poorly invasive/non-metastatic cell lines (Kirschmann et al., 1999). Therefore, the δ EF1 site was analyzed as a candidate regulatory element (Fig. 3A).

3.2. Mechanism of *HP1^{Hsx}* differential regulation

The well-characterized breast cancer cell lines MDA-MB-231, which is highly invasive/metastatic, and MCF-7, which is poorly invasive/non-metastatic, were used to determine the mechanism of differential regulation of *HP1^{Hsx}*. One possible explanation for differential expression of *HP1^{Hsx}* is that a mutation in the *HP1^{Hsx}* gene within MDA-MB-231 cells results in reduced expression. The coding region, splice junctions, portions of the introns, and 150 bp of the promoter region of *HP1^{Hsx}* from MDA-MB-231 and MCF-7 cells were sequenced and compared. No differences between the *HP1^{Hsx}* genomic sequences of the two cell lines were found. Therefore, differential regulation of *HP1^{Hsx}* is not likely due to mutations within the sequenced regions of *HP1^{Hsx}*.

As a second possibility to explain the differential expression, we investigated the DNA methylation status of *HP1^{Hsx}* in both cell types. In many cases transcriptional regulation in cancer cells is under control of DNA methylation, particularly for genes near CpG islands (Dallol et al., 2003). A CpG island within exon one of the *hnRNPA1* gene (–482 to –899 from *HP1^{Hsx}* transcription start) was identified using CpGReport (<http://www.ebi.ac.uk/>). Both strands of the *HP1^{Hsx}* promoter region (bases –900 to +168), including the CpG island, were subjected to bisulfite sequencing to determine DNA methylation of

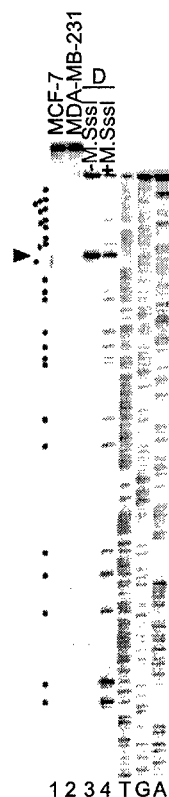


Fig. 2. Absence of detectable 5-methylcytosine (5mC) within the HPI^{Hsx} promoter region in MCF-7 and MDA-MB-231 cells. DNA from both cell lines (lanes 1 and 2) was subjected to a sensitive variation of genomic bisulfite sequencing that is able to detect low levels of 5mC . Analysis of sequences from -220 to $+168$ of the HPI^{Hsx} promoter is shown. Plasmid DNA containing the HPI^{Hsx} promoter methylated *in vitro* by *M. SssI* (lane 4) provides a marker for modified CpG sites (filled circles). Since the plasmid was isolated from a *dcm+* strain of *E. coli*, methylation at a *dcm* site (arrow head) was also detected (lanes 3 and 4). Methylation at the *dcm*- and *M. SssI*-modified sites demonstrates the signal intensity that is commensurate with high and moderate levels of DNA methylation, respectively. Thus, MCF-7 and MDA-MB-231 cells do not have detectable levels of 5mC . Reactions carried out on purified plasmid DNA are labeled D.

HPI^{Hsx} in the MDA-MB-231 and MCF-7 cell lines (Fig. 2 shows the methylation status of bases -220 to $+168$; bases -900 to -220 are not shown). Limited DNA methylation, if any, was observed in either cell type, and the methylation status was unchanged between the two cell lines throughout the HPI^{Hsx} promoter region and exon one, including the CpG island within *hnRNPA1*. Thus, methylation does not appear to be involved in the differential regulation of HPI^{Hsx} .

As a third explanation for the differential regulation, we hypothesized that differential expression might arise through different interactions between transcription factors and *cis*-acting DNA elements of the HPI^{Hsx} promoter in the two cell types. Fragments of the HPI^{Hsx} promoter, including untranslated exon one, were cloned upstream of a *luciferase* reporter gene. These constructs were co-trans-

fected into MDA-MB-231 and MCF-7 cells along with a plasmid containing the *CMV* promoter driving a *lacZ* reporter gene. Expression of luciferase and β -galactosidase was measured in light units. Luciferase expression was normalized to β -galactosidase expression to account for differences in transfection efficiency. Levels of luciferase above background were observed from all constructs, indicating that the cloned promoter region possessed transcriptional activity. A 5.6-fold difference in luciferase activity was observed between the MDA-MB-231 and MCF-7 cell lines for the largest construct, $-600/+143$, containing 13 sequences between the HPI^{Hsx} and the *hnRNPA1* predicted transcription start sites (p -value = 0.0021) (Fig. 3B and C). This difference in expression is similar to that observed for the endogenous HPI^{Hsx} gene between the two cell types (Kirschmann et al., 2000). Deletions that removed successive amounts of 5' sequences showed a reduction in the fold change in expression between the two cell types, suggesting the loss of regulatory sequences (Fig. 3B). In particular, the differences in expression between the $-166/+143$ construct in the two cell types were barely statistically significant (p -value = 0.0608). One caveat of these deletion constructs is that vector DNA sequences are brought into close proximity to the HPI^{Hsx} promoter, possibly contributing to regulation.

To better identify elements involved in differential regulation of HPI^{Hsx} between the two cell types, constructs containing mutations within candidate transcription factor binding sites were analyzed. These constructs allowed for the retention of 600 bp of upstream sequences. Of particular interest was a δ EF1 binding site at position -125 (Fig. 3A). This site can be bound by the human homolog of the chicken δ EF1 protein, AREB6, a transcriptional repressor (Ikeda et al., 1998) that is up-regulated in highly invasive/metastatic cell lines, including MDA-MB-231, compared to poorly invasive/non-metastatic cell lines, such as MCF-7 (Kirschmann et al., 1999). Thus, the δ EF1 binding site was a promising candidate for regulating differential expression of HPI^{Hsx} . A mutation in the conserved δ EF1 site ($m\delta$ EF1) was constructed by replacing the 4 bp core binding site and 5 bp of surrounding sequence with nine cytosines in the context of the $-600/+143$ construct. Differential expression between the two cell lines was still observed (2.87-fold difference, p -value = 0.0004) (Fig. 3C). Therefore, the δ EF1 site does not appear to be involved in differential regulation of HPI^{Hsx} .

Several additional candidate transcription factor-binding sites were also investigated for their effects on differential regulation. These include a c-myc site within exon one of HPI^{Hsx} (Fig. 3A). Mutation of this site (*mm*ycb), replacing the core binding region and surrounding bases with nine cytosines, retained differential expression (nine-fold difference, p -value = 0.0124) (Fig. 3C). Furthermore, mutations of two USF/c-myc sites located at positions -109 and -172 , designated USFd (distal) and USFp (proxi-

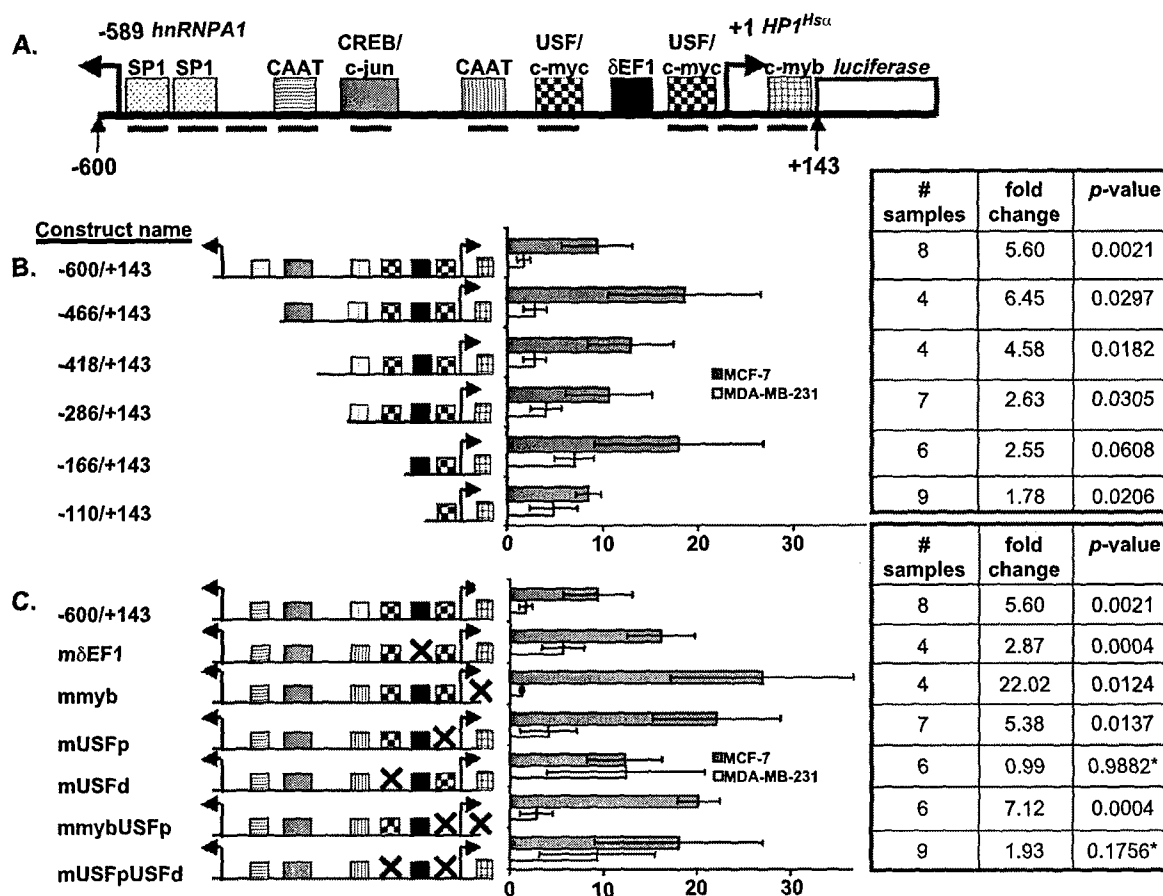


Fig. 3. (A) Diagram of the promoter region between the divergently transcribed *hnRNPA1* and *HP1^{Hsx}* genes. Consensus transcription factor binding sites that were previously shown to be footprinted by protein in HeLa cells (Biamonti et al., 1993) are indicated by black lines. Transcription factor binding sites identified by our bioinformatics searches are indicated by the name above the boxes. (B) Results of luciferase assays from 5' deletion constructs containing *HP1^{Hsx}* promoter fragments fused to a *luciferase* reporter gene. The construct names indicate the 5' and 3' sequence boundaries. The constructs were transfected into a metastatic/highly invasive cell line, MDA-MB-231 and a nonmetastatic/poorly invasive cell line, MCF-7. The number of samples, fold change between expression in MCF-7 and MDA-MB-231 cells, and *p*-values are indicated for each construct. (C) Results of mutational analysis of the *HP1^{Hsx}* promoter constructs. Construct names reflect the site(s) mutated. Asterisks mark *p*-values that show no statistical difference in expression between the two cell lines, indicating a loss of differential regulation.

mal), respectively, were tested for effects on expression (Fig. 3A). Mutations in the individual sites (mUSFp and mUSFd), double mutations of both USF/c-myc sites (mUSFpUSFd), and mutation of USFp in combination with mutation of the c-myc site (mmybUSFp) were examined. All of the mutations replaced the core 4 bp of the transcription factor binding sites and surrounding sequence with eight cytosines. The mUSFp mutation retained differential expression (5.38-fold difference, *p*-value=0.0137) (Fig. 3C). The mmybUSFp double mutant also retained differential expression (7.12-fold difference, *p*-value=0.0004). In contrast, mUSFd eliminated differential expression (0.99-fold difference, *p*-value=0.9882). Consistent with this finding, mUSFpUSFd disrupted differential expression (*p*-value=0.1756). Thus, the USFd site at position -172 is required for differential expression of *HP1^{Hsx}* between MCF-7 and MDA-MB-231 cell lines.

3.3. Human *HP1^{Hsx}* functions similar to *Drosophila* *HP1*

In addition to understanding *HP1^{Hsx}* regulation, we also aimed to understand the functional properties of this protein. *HP1* proteins are highly conserved between species, suggesting related functions. *HP1* was first identified in *Drosophila* where functional studies have been performed, demonstrating a role in gene silencing (Weiler and Wakimoto, 1995). Three human *HP1* family members, *HP1^{Hsx}*, *HP1^{Hsβ}* and *HP1^{Hsγ}*, show a high degree of amino acid sequence identity with *Drosophila* *HP1*; however, it is unclear which family member is the functional homolog of *Drosophila* *HP1*. *HP1^{Hsγ}* shows 44% overall amino acid sequence identity with *Drosophila* *HP1*, 60% in the CD, and 38% in the CSD. These values are similar to those of *HP1^{Hsx}* showing 44% overall, 65% in the CD, and 43% in the CSD. *HP1^{Hsβ}* has slightly greater overall amino acid identity to *Drosophila* *HP1* than *HP1^{Hsx}* (46% vs. 44%).

The HP1^{Hsβ} CD shows slightly more identity with the *Drosophila* HP1 CD than the HP1^{Hsγ} and HP1^{Hsα} CDs (68% vs. 60% and 65%, respectively). In contrast, the HP1^{Hsβ} CSD shows less identity to the *Drosophila* HP1 CSD than the HP1^{Hsα} CSD (39% vs. 43%). In sum, comparisons of the amino acid sequences of human and *Drosophila* HP1 identified only minor differences in the percent identity without immediately suggesting a functional homologue.

Another protein feature that might suggest similar function between *Drosophila* HP1 and a human HP1 is the chromosome localization pattern. *Drosophila* HP1 shows enrichment at heterochromatic regions and localizes to approximately 200 euchromatic sites on larval polytene chromosomes (Fanti et al., 2003). HP1^{Hsα} and HP1^{Hsβ} predominantly localize to centric heterochromatin, showing partial overlap with anti-centromere antibodies (Minc et al., 1999). In contrast, HP1^{Hsγ} localizes to centric heterochromatin and euchromatic regions (Minc et al., 2000). Based on this localization data, HP1^{Hsγ} appears to have a more similar pattern to that of *Drosophila* HP1.

To investigate the functional properties of the HP1^{Hsα} protein and determine whether it is a functional homologue of *Drosophila* HP1, we generated transgenic *Drosophila* that expressed an HP1^{Hsα}-EGFP fusion gene under the control of an *hsp70* heat shock promoter. Homozygous HP1^{Hsα}-EGFP larvae were heat shocked 1 h at 37°C and allowed

to recover for 2 h at room temperature. Salivary glands were dissected from the larvae, squashed and stained with antibodies that recognize EGFP and *Drosophila* HP1. The results indicated that HP1^{Hsα}-EGFP localized to the chromocenter (the site of fusion of all the centromeres), the heterochromatic fourth chromosome, and euchromatic sites in a pattern that completely overlapped with endogenous *Drosophila* HP1 (Fig. 4A). It was possible that co-localization was due to interactions between the HP1^{Hsα} CSD and the *Drosophila* HP1 CSD, since CSDs have been shown to dimerize (Cowieson et al., 2000). Therefore, we assayed the localization of HP1^{Hsα}-EGFP in larvae lacking endogenous *Drosophila* HP1. HP1^{Hsα}-EGFP showed the same pattern of localization on larval polytenes with or without endogenous HP1, indicating that HP1^{Hsα} associates with chromosomes by similar mechanisms as *Drosophila* HP1 (Fig. 4B).

Overexpression of *Drosophila* HP1 enhances silencing of genes repressed by heterochromatin (Weiler and Wakimoto, 1995). To test whether HP1^{Hsα}-EGFP has a similar function, HP1^{Hsα}-EGFP was overexpressed (two-fold over endogenous HP1) in stocks carrying a tagged *hsp26* heat shock gene inserted at different heterochromatic locations. Expression of HP1^{Hsα}-EGFP by daily heat shock resulted in a 40% reduction in *hsp26* expression from a centric transgene (Fig. 5). Similarly, expression of HP1^{Hsα}-EGFP resulted in a 50% reduction in *hsp26* expression from a telomeric transgene

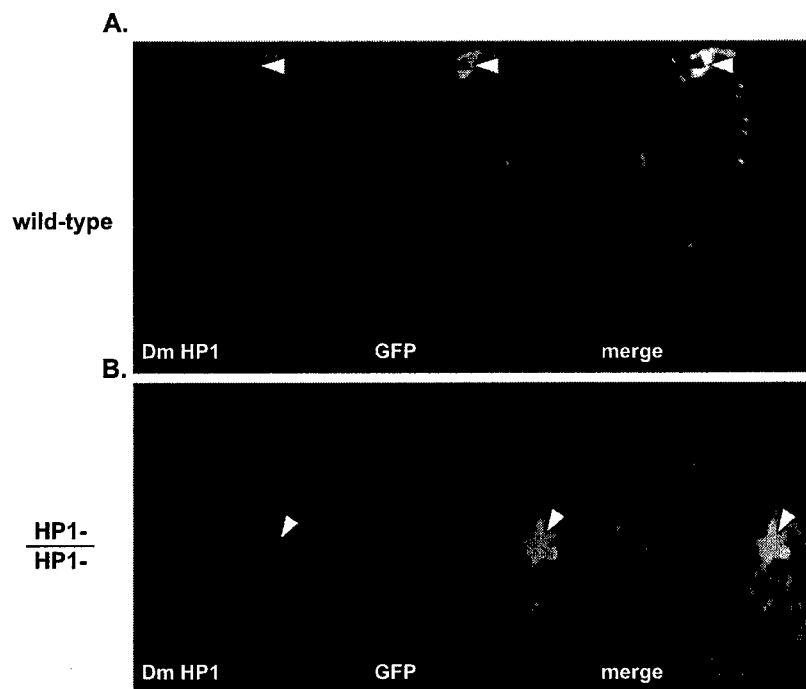


Fig. 4. (A) Polytene chromosomes from third instar larvae salivary glands from a stock carrying an HP1^{Hsα}-EGFP transgene. The chromosomes were fixed, squashed and stained with antibodies against *Drosophila* HP1 (DmHP1; red) and GFP (green). The DmHP1 antibody does not recognize HP1^{Hsα} by Western blot analysis (data not shown). Co-localization is observed in yellow. (B) Polytene chromosomes from HP1 mutants containing the HP1^{Hsα}-EGFP transgene stained with DmHP1 (red) and GFP (green). Arrowheads denote the chromocenter (formed by the fusion of all centromeres).

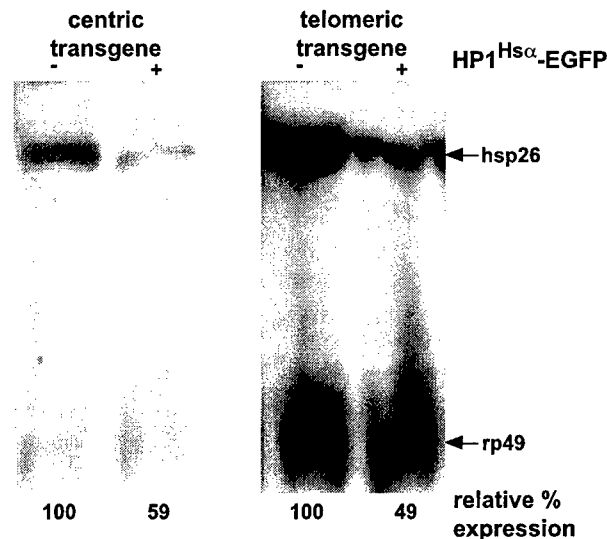


Fig. 5. Effects of $HP1^{Hs\alpha}$ -EGFP expression on gene silencing. Flies expressing the $HP1^{Hs\alpha}$ -EGFP transgene were crossed to flies that contain a reporter *hsp26* transgene inserted at a heterochromatic or telomeric location. RNA was isolated from the heat-shocked progeny and analyzed by northern analysis with radiolabeled sequences corresponding to the *hsp26* reporter gene and the *rp49* loading control.

(Fig. 5). These data demonstrate that the human $HP1^{Hs\alpha}$ protein participates in gene silencing and has similar functions as Drosophila HP1.

Several homozygous lethal mutations exist in *Su(var)2-5*, the Drosophila gene encoding HP1 (Eissenberg and Hartnett, 1993). These mutations were used to determine whether $HP1^{Hs\alpha}$ -EGFP could rescue lethality. Flies carrying the $HP1^{Hs\alpha}$ -EGFP transgene that were heterozygous for a mutant allele of the gene encoding HP1 [*Su(var)2-5⁰⁴*] balanced over a chromosome possessing a *Curly wing* mutation were crossed to flies heterozygous for a second mutant allele of the gene encoding HP1 [*Su(var)2-5⁰²*] over the *Curly wing* balancer chromosome. From this cross, *Curly wing* homozygotes, 25% of total progeny, die as early embryos. Flies heterozygous for the *Curly wing* balancer chromosome and a *Su(var)2-5* allele, 50% of the total progeny, were viable. The final class of progeny, heteroallelic for the *Su(var)2-5* mutant alleles, constituting 25% of the total progeny, would be lethal if no rescue is observed. Rescue of such individuals would give rise to straight wing adults. When $HP1^{Hs\alpha}$ -EGFP was expressed by daily heat shock treatment, 2% (4/218 adults) of the total progeny had straight wings, suggesting partial function of the $HP1^{Hs\alpha}$ -EGFP fusion protein. More convincing, 14% (35/244 adults) of total progeny were rescued to adulthood by expressing $HP1^{Hs\alpha}$ without the EGFP tag. These data suggest that despite exhibiting the correct pattern of localization and gene silencing effects, EGFP was limiting the function of $HP1^{Hs\alpha}$. The ability of the untagged version to show appreciable rescue (14% vs. 25% for complete rescue) strongly suggests that $HP1^{Hs\alpha}$ is a functional homolog of the Drosophila HP1 protein.

4. Discussion

4.1. Conservation of gene structure

The *HP1* genomic structure is conserved from Drosophila to humans. Drosophila *Su(var)2-5*, mouse *mHP1 α* , *M31*, and *M32*, and human *HP1^{Hs α}* , *HP1^{Hs β}* , and *HP1^{Hs γ}* are each comprised of five exons and four introns. Translation start begins in exon two in Drosophila *Su(var)2-5*, mouse *mHP1 α* and *M31*, and human *HP1^{Hs α}* , *HP1^{Hs β}* and *HP1^{Hs γ}* . In contrast, the translation start of *M32* is within exon three. Due to the insertion of an intron within exon one, *M32* exon three corresponds to that of exon two in the other HP1 genes (Jones et al., 2001). The CD of all of the mammalian HP1 genes, except for *M32*, is contained in exons two and three (Jones et al., 2001). Exons three and four of *M32* have fused to become exon four. Therefore, the CD of *M32* is found within exons three and four. The CSD of all of the mammalian HP1 genes is found in exons four and five (Jones et al., 2001). The amino acids at the splice junctions are conserved in all the mammalian HP1 family members, except *M32*, but are distinctly different for Drosophila *Su(var)2-5*. Therefore, the genomic structure, but not the intron/exon boundaries, are conserved from Drosophila *Su(var)2-5* to human HP1 family members.

The *Su(var)2-5* gene, at cytological position 28F1-2, encodes a protein that is commonly referred to as Drosophila HP1 and sometimes referred to as HP1a (Smothers and Henikoff, 2001). There are two additional HP1-like genes located at cytological positions 87C and 94C4, called *HP1b* and *HP1c*, respectively. The proteins encoded by these genes do not exhibit a chromosomal distribution pattern that significantly overlaps with HP1 (Smothers and Henikoff, 2001). In addition, these two proteins have limited amino acid sequence identity with HP1; *HP1b* is 44% identical and *HP1c* is 31% identical to HP1. Furthermore, *HP1b* and *HP1c* do not have a conserved genomic structure (<http://www.flybase.org/>) with the mammalian members of the HP1 family, as does Drosophila HP1. Thus, HP1 was exclusively used for comparisons in this study.

The three human genes encoding *HP1^{Hs α}* , *HP1^{Hs β}* and *HP1^{Hs γ}* located at 12q13, 17q21 and 7q15, respectively, have several corresponding pseudogenes present in the human genome. In our screen of the human genomic library, we recovered one processed *HP1^{Hs α}* pseudogene (data not shown) corresponding to sequences on chromosome 3p11.1 (<http://www.ncbi.nlm.nih.gov/>). Five pseudogenes have been previously published for *HP1^{Hs β}* , and four for *HP1^{Hs γ}* (Jones et al., 2001). The pseudogenes of *HP1^{Hs β}* , containing few, if any, introns are found on chromosomes 1q32, 3q26, 14q24, Xp22, and Xq11 (<http://www.ncbi.nlm.nih.gov/>) suggesting they are processed pseudogenes. Eleven *HP1^{Hs γ}* pseudogenes are found on NCBI Genbank, four of which correspond to the previously published pseudogenes (<http://www.ncbi.nlm.nih.gov/>; Jones et al., 2001). The pseudogenes of *HP1^{Hs γ}* , containing few, if any, introns

are found on chromosomes 2q24, 3p22, 5q22, 6q22.2, 11p11, 11p14, 11q14, 12p13, 12q23, 16p13, and 18p11 (<http://www.ncbi.nlm.nih.gov/>). Thus, each of the functional human HP1 family members is encoded by separate unlinked genes that have multiple pseudogenes scattered throughout the genome.

4.2. HP1^{Hsα} regulation in breast cancer metastasis

HP1^{Hsα} is down-regulated in highly invasive/metastatic breast cancer cells compared to poorly invasive/non-metastatic cells (Kirschmann et al., 2000). This regulation likely occurs, at least in part, at the transcriptional level and does not involve differential DNA methylation. An analysis of the DNA sequences in the HP1^{Hsα} promoter region identified potential binding sites for transcriptional regulators that might be involved in differential regulation. Only three of the binding sites shown in Fig. 3, the two SP1 sites immediately upstream of *hnRNPA1*, the CAAT box (position –244) and the proximal USF/c-myc site (position –109), are conserved between mouse and humans; none of the elements can be identified upstream of the gene encoding Drosophila HP1.

Mutation of a δ EF1 binding site (at position –125), which associates with the AREB6 protein found to be up-regulated in highly invasive/metastatic breast cancer cells (Kirschmann et al., 1999), does not appear to be involved in the differential expression of HP1^{Hsα}. Mutation of a c-myc binding site at position +97 and a USF/c-myc site at position –109 does not appear to play a role in differential regulation either. In contrast, mutation of a USF/c-myc site at position –172 abolishes differential regulation. This USF/c-myc was protected from DNase I digestion, suggesting occupancy by protein(s) in HeLa cells (Biamonti et al., 1993). USF/c-myc sites, commonly called E-boxes, are frequently bound by a variety of proteins, including USF and Myc family members. USF proteins are involved in both gene silencing and activation, sometimes at the same site under different circumstances (Goueli and Janknecht, 2003). Myc proteins are also involved in both gene activation and repression, depending on their dimerization partner (Queva et al., 1998). Therefore, the function of the distal USF/c-myc site in the differential regulation of HP1^{Hsα} is difficult to predict and warrants further investigation.

4.3. Function of HP1^{Hsα}

Our data strongly suggest that HP1^{Hsα} is a functional homolog of the Drosophila HP1 protein. The results showing that HP1^{Hsα} can localize to the same sites on polytene chromosomes as Drosophila HP1 are consistent with previously published results (Ma et al., 2001). We extend these findings by demonstrating that HP1^{Hsα} exhibits the Drosophila HP1 pattern of localization even in the absence of Drosophila HP1. These results suggest a conserved mechanism for localization. Previously pub-

lished results show that HP1^{Hsα} can enhance silencing induced by a transgene array in Drosophila (Ma et al., 2001). These arrays have similar, yet distinctly different, properties than heterochromatin (Prasad-Sinha et al., 2000). Our results clearly demonstrate that HP1^{Hsα} can participate in heterochromatin formation and silence euchromatic genes placed within heterochromatin. Thus, HP1^{Hsα} has gene silencing functions similar to Drosophila HP1.

Species specificity of protein function can be addressed by determining whether a protein from one organism can provide complete function of the homologous protein in another organism. HP1^{Hsα} can rescue the lethality of a Drosophila HP1 homozygous mutant; therefore, HP1^{Hsα} is a functional homolog of Drosophila HP1. In contrast to our findings, the mouse M31 protein was unable to rescue mutant phenotypes associated with *S. pombe* Swi6 mutants (Wang et al., 2000). Rescue was obtained, however, when the Swi6 CSD was substituted for the M31 CSD (Wang et al., 2000). The overall amino acid sequence identity between *S. pombe* Swi6 and mouse M31 is 37%. This is much less than the 44% overall amino acid sequence identity between Drosophila HP1 and human HP1^{Hsα}. In particular, the amino acid sequence identity between the *S. pombe* Swi6 CSD and the mouse M31 CSD is 39%, whereas the amino acid sequence identity between Drosophila HP1 CSD and the human HP1^{Hsα} CSD is 43%. Therefore, the CSD of Drosophila HP1 and human HP1^{Hsα} is more conserved than the CSD of *S. pombe* Swi6 and mouse M31. The amino acid sequence differences between mouse and *S. pombe* might explain the species-specificity observed.

4.4. Model for HP1^{Hsα} function in breast cancer metastasis

Given the conserved function of HP1^{Hsα} in gene regulation, one possible role for HP1^{Hsα} in breast cancer metastasis is gene silencing. Accordingly, the HP1^{Hsα} gene would be expressed in normal and primary breast cancer tumor cells where it produces protein that functions to silence genes required for invasion and metastasis. In highly invasive/metastatic breast cancer cells, HP1^{Hsα} expression is reduced and less HP1^{Hsα} protein is available to carry out gene silencing functions. Clearly in Drosophila and mice HP1 affects gene expression in a dosage-dependent manner (Weiler and Wakimoto, 1995; Festenstein et al., 1999). According to this model, loss of silencing would occur at genes encoding proteins that are required for invasion and metastasis. Therefore, the identification of HP1^{Hsα} regulated genes is a goal for future investigation.

Acknowledgements

Funding for the breast cancer research was supported by the DOD Breast Cancer Research Program (DAMD17-02-1-

0424). The *Drosophila* research was supported by NIH (GM61513). We thank members of the Wallrath lab for comments regarding the manuscript.

References

- Biamonti, G., Bassi, M.T., Cartegni, L., Mechta, F., Buvoli, M., Cobiainchi, F., Riva, S., 1993. Human hnRNP protein A1 gene expression. Structural and functional characterization of the promoter. *J. Mol. Biol.* 230, 77–89.
- Brehm, A., Tufeland, K.R., Aasland, R., Becker, P.B., 2004. The many colours of chromodomains. *Bioessays* 26, 133–140.
- Cowieson, N.P., Partridge, J.F., Allshire, R.C., McLaughlin, P.J., 2000. Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr. Biol.* 10, 517–525.
- Dallol, A., Krex, D., Hesson, L., Eng, C., Maher, E.R., Latif, F., 2003. Frequent epigenetic inactivation of the SLIT2 gene in gliomas. *Oncogene* 22, 4611–4616.
- Eissenberg, J.C., Elgin, S.C., 2000. The HP1 protein family: getting a grip on chromatin. *Curr. Opin. Genet. Dev.* 10, 204–210.
- Eissenberg, J.C., Hartnett, T., 1993. A heat shock-activated cDNA rescues the recessive lethality of mutations in the heterochromatin-associated protein HP1 of *Drosophila melanogaster*. *Mol. Gen. Genet.* 240, 333–338.
- Ekwall, K., Javerzat, J.P., Lorentz, A., Schmidt, H., Cranston, G., Allshire, R., 1995. The chromodomain protein Swi6: a key component at fission yeast centromeres. *Science* 269, 1429–1431.
- Fanti, L., Berloco, M., Piacentini, L., Pimpinelli, S., 2003. Chromosomal distribution of heterochromatin protein 1 (HP1) in *Drosophila*: a cytological map of euchromatic HP1 binding sites. *Genetica* 117, 135–147.
- Festenstein, R., Sharghi-Namini, S., Fox, M., Roderick, K., Tolaini, M., Norton, T., Saveliev, A., Kioussis, D., Singh, P., 1999. Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner. *Nat. Genet.* 23, 457–461.
- Goueli, B.S., Janknecht, R., 2003. Regulation of telomerase reverse transcriptase gene activity by upstream stimulatory factor. *Oncogene* 22, 6098–6103.
- Gum Jr., J.R., Hicks, J.W., Crawley, S.C., Dahl, C.M., Yang, S.C., Robertson, A.M., Kim, Y.S., 2003. The MUC3A human intestinal mucin: Initiation of transcription from a TATA-less promoter and comparison to the MUC3B amino terminus. *J. Biol. Chem.*
- Ikeda, K., Halle, J.P., Stelzer, G., Meisterernst, M., Kawakami, K., 1998. Involvement of negative cofactor NC2 in active repression by zinc finger-homeodomain transcription factor AREB6. *Mol. Cell Biol.* 18, 10–18.
- Jones, D.O., Mattei, M.G., Horsley, D., Cowell, I.G., Singh, P.B., 2001. The gene and pseudogenes of Cbx3/mHP1 gamma. *DNA Seq.* 12, 147–160.
- Kirschmann, D.A., Seftor, E.A., Nieva, D.R., Mariano, E.A., Hendrix, M.J., 1999. Differentially expressed genes associated with the metastatic phenotype in breast cancer. *Breast Cancer Res. Treat.* 55, 127–136.
- Kirschmann, D.A., Lininger, R.A., Gardner, L.M., Seftor, E.A., Otero, V.A., Ainsztein, A.M., Earnshaw, W.C., Wallrath, L.L., Hendrix, M.J., 2000. Down-regulation of HP1Hsalpha expression is associated with the metastatic phenotype in breast cancer. *Cancer Res.* 60, 3359–3363.
- Kladde, M.P., Xu, M., Simpson, R.T., 1996. Direct study of DNA-protein interactions in repressed and active chromatin in living cells. *EMBO J.* 15, 6290–6300.
- Ma, J., Hwang, K.K., Worman, H.J., Courvalin, J.C., Eissenberg, J.C., 2001. Expression and functional analysis of three isoforms of human heterochromatin-associated protein HP1 in *Drosophila*. *Chromosoma* 109, 536–544.
- Minc, E., Allory, Y., Worman, H.J., Courvalin, J.C., Buendia, B., 1999. Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. *Chromosoma* 108, 220–234.
- Minc, E., Courvalin, J.C., Buendia, B., 2000. HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet. Cell Genet.* 90, 279–284.
- Muchardt, C., Guilleme, M., Seeler, J.S., Trouche, D., Dejean, A., Yaniv, M., 2002. Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep.* 3, 975–981.
- Nielsen, A.L., Oulad-Abdelghani, M., Ortiz, J.A., Remboutsika, E., Chambon, P., Losson, R., 2001. Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol. Cell* 7, 729–739.
- Platero, J.S., Hartnett, T., Eissenberg, J.C., 1995. Functional analysis of the chromo domain of HP1. *EMBO J.* 14, 3977–3986.
- Prasad-Sinha, J., Furuyama, T., Dorer, D., Harte, P.J., 2000. The Polycomb Group proteins ESC and E(Z) are required for repeat-induced gene silencing. 41st Annual *Drosophila* Research Conference. The Genetics Society of America, Pittsburgh, PA, p. 231B.
- Queva, C., Hurlin, P.J., Foley, K.P., Eisenman, R.N., 1998. Sequential expression of the MAD family of transcriptional repressors during differentiation and development. *Oncogene* 16, 967–977.
- Shaffer, C.D., Wuller, J.M., Elgin, S.C., 1994. Raising large quantities of *Drosophila* for biochemical experiments. *Methods Cell Biol.* 44, 99–108.
- Smothers, J.F., Henikoff, S., 2000. The HP1 chromo shadow domain binds a consensus peptide pentamer. *Curr. Biol.* 10, 27–30.
- Smothers, J.F., Henikoff, S., 2001. The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. *Mol. Cell Biol.* 21, 2555–2569.
- Wallrath, L.L., Burnett, J.B., Friedman, T.B., 1990. Molecular characterization of the *Drosophila melanogaster* urate oxidase gene, an ecdysone-repressible gene expressed only in the malpighian tubules. *Mol. Cell Biol.* 10, 5114–5127.
- Wang, G., Ma, A., Chow, C.M., Horsley, D., Brown, N.R., Cowell, I.G., Singh, P.B., 2000. Conservation of heterochromatin protein 1 function. *Mol. Cell Biol.* 20, 6970–6983.
- Weiler, K.S., Wakimoto, B.T., 1995. Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genet.* 29, 577–605.

**Deciphering the code of silence: Mechanisms of gene repression with
connections to human disease**

Laura E. Norwood, Karrie A. Hines, and Lori L. Wallrath*

Department of Biochemistry, University of Iowa, Iowa City, Iowa, 52242

Address for Correspondence (*):

Lori L. Wallrath

The Department of Biochemistry

3136 MERF

University of Iowa

Iowa City, Iowa 52242

Phone: 319-335-7920

Fax: 319-384-4770

E-mail: lori-wallrath@uiowa.edu

Key words: gene silencing, heterochromatin, Heterochromatin Protein 1, histone code,
ncRNA, Polycomb, position effect variegation, RNAi

Abstract

Proper regulation of gene expression is required for normal growth and development. Developmental abnormalities and disease can result from the misregulation of gene expression. While some genes are clearly controlled by mechanisms of activation, a significant proportion of genes are controlled by mechanisms of silencing. Silencing mechanisms can account for the inactivation of the approximately one thousand genes on the human X-chromosome, for uniparental inactivation of chromosomal domains, and for inactivation of individual genes. Common features of silencing events include initiation by RNAs, either as non-coding RNAs expressed from a single gene that remain associated near the site of synthesis or as small interfering RNA molecules that act in *trans*. Other common epigenetic features include specific histone modifications, which generate a code for silent chromatin, and DNA methylation. Non-histone chromosomal proteins have been discovered that associate with specific histone modifications. These non-histone chromosomal proteins, “translators” of the histone code, include Heterochromatin Protein 1 and Polycomb, two factors that play a role in organizing chromatin structure. This review discusses discoveries made on gene silencing systems and their connections with disease.

Gene silencing phenomena

Expression of genes within the human genome, estimated to contain approximately 25,000 genes, requires proper regulation. It would be deleterious for a cell to express its full complement of genes at once. Therefore, mechanisms to differentially activate and silence genes are used to coordinate gene expression on the basis of developmental timing and/or cell-type specificity. Mechanisms of gene activation have been well studied¹, however, mechanisms of gene repression, sometimes referred to as gene silencing, are just beginning to be understood. This review discusses classical cases of gene silencing as well as modern molecular discoveries that have revealed mechanisms of gene silencing. As anticipated from early studies of genome packaging, chromatin-associated proteins play a central role in the process of gene regulation, where an epigenetic “code” for silencing has emerged. Unanticipated newcomers involved in gene silencing include non-coding and small interfering RNA molecules. Taken together, the molecular basis for gene silencing is beginning to be deciphered, yet much remains to be understood.

Position effect variegation

In the 1920's and 1930's, classical genetic analysis was carried out by subjecting *Drosophila* (fruit flies) to X-rays. Through such early mutagenesis experiments, Muller identified a special class of mutant flies called “eversporting” that possessed red and white mosaic eyes, with red being the wild type color². Later, cytological and molecular data showed that Muller's eversporting mutant phenotype was due to chromosomal translocations (relocation of chromosomal fragments to a new location in the genome) of

the *white* gene, which encodes a protein required to generate red eyes³. Translocation of the *white* gene near a centromere results in silencing in a subset of cells (Figure 1A). This gene silencing phenomenon was designated position effect variegation (PEV)³ because it was caused by the new position of the gene in the genome, rather than a mutation in the sequence of the gene.

Following the discovery of PEV, several genetic screens were performed in *Drosophila* to identify dominant mutations that enhance [*Enhancer of variegation*, *E(var)s*] or suppress [*Suppressors of variegation*, *Su(var)s*] PEV (Figure 1A)^{4,5}. Some *Su(var)* loci were found to be haplo-suppressors and triplo-enhancers of PEV⁵. Loss of one copy of the *Su(var)* gene suppresses variegation and addition of an extra copy of the *Su(var)* gene enhances variegation, indicating a dosage sensitivity of the variegating gene for the products of these *Su(var)* genes. Such findings suggest that critical levels of modifier proteins are essential for proper gene expression. Many of the *E(var)* and *Su(var)* genes have been found to encode chromatin proteins and enzymes that modify other chromatin proteins, suggesting a role in gene regulation.

PEV is not restricted to *Drosophila*, as examples have been observed in yeast and mice. In yeast, genes inserted near centromeres, telomeres, and the silent mating type locus exhibit variegated expression⁶⁻⁸. PEV is best demonstrated by red and white sectoring in yeast colonies that occur when an *ade6* transgene inserted near centric regions is silenced, leading to red coloration within the yeast colony² (Figure 1B). Likewise, in mice, transgenes inserted near centromeres show PEV. This has been demonstrated for a *CD2*

transgene lacking a regulatory region that protects it from genomic position effects^{9, 10}.

The *CD2* transgene encodes a transmembrane receptor that is expressed on a cell autonomous basis. When *CD2* is inserted near a centromere, PEV is observed as a subpopulation of cells lacking CD2 receptors on their surface⁹ (Figure 1C). Thus, genomic regions such as those surrounding centromeres in diverse species share properties that give rise to PEV.

X-chromosome dosage compensation

In many species, males and females have different numbers of sex chromosomes. In order to compensate for differential expression of genes on sex chromosomes, processes of dosage compensation have evolved¹¹. Dosage compensation is facilitated by different mechanisms for many of the classical model organisms where X-chromosome number varies between males and females. In insects, such as *Drosophila*, genes on the X-chromosome in males are expressed at two-fold greater levels than from the female X¹¹. In the worm *Caenorhabditis elegans*, genes on both female X-chromosomes are down-regulated to half the expression level of the single male X¹¹. A different strategy is taken in mammals, where genes on one of the female X-chromosomes are inactivated, resulting in expression from the one active X-chromosome equaling that observed from the single male X-chromosome¹¹. In all cases, the net result is to balance expression of genes on the X-chromosome between the sexes, although the means to achieve this balance occurs by different mechanisms. The species-specific strategies to achieve this balance have something in common: they all utilize aspects of chromatin packaging to change the expression of X-linked genes at the level of transcription.

From cytological analyses of the mammalian inactive X (Xi), also known as the Barr body, alterations in chromatin packaging are clearly associated with gene silencing^{12, 13}. The Xi, which becomes inactivated in the embryo before implantation of the blastocyst, forms a compact, mitotically stable structure localizing to the nuclear periphery^{11, 13}. In some mammals, such as marsupials, the paternal X-chromosome is always designated as the chromosome to undergo Barr body formation¹¹. In placental mammals, such as cats and humans, the paternal X-chromosome is always inactivated in extraembryonic tissues, but within the embryonic tissue the Xi is randomly chosen between the paternally and maternally inherited X-chromosomes¹¹. Random inactivation can lead to the mosaic expression of X-linked alleles within an individual, hence the molecular basis of the female calico cat¹⁴.

Mosaicism at the individual level occurs when the paternal X-chromosome is inactivated in a fraction of the cells and the maternal X-chromosome is inactivated in the remainder of the cells, leading to different cells within an individual expressing different alleles of the same gene¹⁵. X-inactivation can have implications for human disease. For example, Burn *et al* (1986) reported monozygotic twin girls, only one of which displayed phenotypes associated with the X-linked disease Duchenne Muscular Dystrophy (DMD)¹⁶. In the affected twin, the paternally inherited X-chromosome was inactivated in the majority of cells¹⁶. In contrast, in the unaffected twin the maternally inherited X-chromosome was inactivated in the majority of cells¹⁶. These data strongly imply that the mutant allele causing DMD was maternally inherited and that non-random X-inactivation

resulted in DMD in the affected twin. This phenomenon is called “skewing” and involves non-random X-chromosome inactivation within an individual. Usually skewing contributes to a lower percentage of females with X-linked diseases¹⁵.

X-chromosome inactivation typically involves the majority of genes on the chromosome, with the exception of a class known as “escapers”. For the most part, escaper genes neither share common expression patterns nor encode proteins with related functions. Their chromosomal positions cluster, though, with the majority mapping to the small arm of the X-chromosome, suggesting chromosome domain organization¹⁴. Whether these genes escape inactivation or are reactivated following initial gene silencing is not clear. Recent data favor a mechanism whereby escaper genes lack the ability to maintain silencing¹⁴.

Genomic imprinting

Imprinting is the process whereby a gene on an autosomal chromosome has a different expression pattern depending on whether it is inherited maternally or paternally. This mechanism of gene regulation is found in both animals and plants¹⁷. In contrast to X-chromosome inactivation that is nearly chromosome-wide, autosomal imprinting only affects a small number of genes per chromosome¹⁸. These imprinted genes are frequently arranged in clusters, such as in the Beckwith-Wiedemann syndrome (BWS) cluster that spans one Mb of DNA on chromosome 11 and contains 15 genes, nine of which are imprinted^{18, 19}.

In mammals, improper expression or mutation of imprinted genes generally leads to growth defects, impairment of brain function, and cancer. For example, BWS in humans is associated with alterations in the expression of the two imprinted genes, *IGF2* and *CDKN1C* within the BWS cluster on chromosome eleven¹⁹. Misregulation of genes within this cluster lead to developmental defects and a predisposition to Wilm's tumor¹⁹,²⁰. *IGF2* is normally expressed from the paternally inherited chromosome; *CDKN1C* is normally expressed from the maternally inherited chromosome¹⁹. Uniparental inheritance of the BWS region from the male, chromosomal translocations that remove an allele of either imprinted gene, or point mutations within the maternal *CDKN1C* gene that disturb its expression all lead to disruption of the imprinted pattern, resulting in BWS¹⁹.

Chromatin structure regulates gene expression

Silencing genes through PEV, X-inactivation, imprinting and other phenomena is directly linked to chromatin packaging. In eukaryotic cells, the genome is packaged into higher order chromatin structures, with the smallest unit of chromatin packaging being the nucleosome. The nucleosome consists of 146 ± 2 bp of DNA (1.7 turns of the DNA superhelix) wrapped around an octamer of histones (a tetramer of the histones H3/H4 and two dimers of the histones H2A/H2B dimer)^{21,22}. Nucleosomes wrap the DNA, leaving 20-80 bp of double stranded DNA to serve as a linker between nucleosomes; the linker size varies according to the organism and genomic context²³. The interaction between DNA and the core histones results in a five-fold compaction of the DNA²⁴. Nucleosome placement with respect to gene sequences is critical for determining whether a gene will

be expressed. Nucleosomes positioned over TATA boxes and/or regulatory elements can block access of RNA polymerase and result in gene repression²⁵. In some cases, chromatin-remodeling machines are necessary to slide or displace nucleosomes from genes prior to activation²⁶.

Further condensation of nucleosome-associated DNA is facilitated by histone H1, which binds linker DNA, compacting chromatin into a 30 nm fiber²⁴. In general, association of H1 correlates with gene repression²⁷. Packaging beyond the 30 nm fiber requires additional non-histone chromosomal proteins²³. How these proteins generate higher order chromatin structures, such as those observed near centromeres, is a topic under investigation.

In most eukaryotic organisms the genome is packaged into two general types of chromatin: euchromatin and heterochromatin. Euchromatin contains the majority of genes and is packaged into irregular nucleosome arrays. Nucleosome-free regions are detected as sites hypersensitive to nucleases and usually map to the regulatory regions of genes²⁸. In contrast, heterochromatin contains large quantities of repetitive elements in the genome and is packaged into very regular nucleosome arrays²⁸. Heterochromatic regions are typically located near centromeres and telomeres and remain as condensed chromatin throughout the cell cycle. With these data in mind, it is understandable how placement of a euchromatic gene into heterochromatin alters gene expression, as observed for PEV. The translocated gene appears to take on the packaging state of the new genomic environment.

Epigenetic mechanisms of gene regulation

Euchromatin and heterochromatin are not only distinguished by DNA sequence content and nucleosome array organization, but also by modifications of amino acid residues present in histone proteins. In particular, the amino terminal tails of histones that extend from the nucleosome core are post-translationally modified in numerous ways. Particular modifications have been shown to correlate with specific functions (Table 1)^{29,30}. These modifications serve as epigenetic marks that can be perpetuated through cell divisions³¹. One hypothesis states that these modifications represent a “histone code”, serving as docking sites for specific proteins that regulate chromatin structure, gene expression, entry into mitosis and additional functions³¹. A second hypothesis suggests that the histone modifications regulate the levels of compaction and define large-scale domains on chromosomes³². These hypotheses are not mutually exclusive and suggest multiple mechanisms to control gene expression.

Extensive analysis has been carried out on the modifications that occur at lysine residues within the histone H3 and H4 amino-terminal tails. Active transcription is typically associated with acetylation of lysines 9 and 14, and methylation of lysine 4 of histone H3 tails²⁹. These modifications are generated by histone acetyltransferases (HATs) and histone methyltransferases, respectively. In contrast, gene silencing is frequently associated with hypoacetylation of histone H3 and H4 tails, as well as methylation of lysine 9 of histone H3³¹. In accordance with these “codes”, euchromatin and

heterochromatin have distinct histone modifications corresponding to their relative transcriptional activity.

Connections have been made between histone modifications and another epigenetic regulator of gene expression, DNA methylation. Cytosine residues within the genome can be methylated by DNA methyltransferases. This modification is typically associated with gene silencing. The mechanism of silencing involves the recruitment of methyl-DNA binding proteins to the methylated cytosines, which in turn recruit histone deacetylases (HDACs)³³. Therefore, sites of DNA methylation can direct patterns of histone deacetylation throughout the genome.

Similar to the targeting of histone deacetylation, DNA methylation has also been linked with histone methylation in *Neurospora*. A genetic screen performed to identify mutations that disrupt DNA methylation identified a histone methyltransferase³⁴. Adding further support for this connection, the maintenance DNA methyltransferase mutant (*met1*) in *Arabidopsis* shows a reduction in the amount of histone H3 lysine 9 methylation³⁵. Taken together, these data indicate that DNA methylation can be dependent on histone methylation and vice versa. Such findings invoke a multi-tiered regulation system, with at least two components (histone and DNA methylation) that can be mitotically passed on to ensure maintenance of the silent state.

Methylation is one of the most stable covalent histone modifications identified and is thought to serve as a “memory” for silent chromatin³⁶. No histone demethylase has been

discovered to date³⁶. Instead, removal of the histone mark may occur through active histone replacement, cleavage of the histone tail, or dilution of the histone methylation through successive rounds of DNA replication and chromatin packaging³⁷. Acetylation, on the other hand, is a highly dynamic histone modification. HATs and HDACs alter the state of acetylation of histones, leading to transcriptional activation or repression, respectively³⁷. Such features allow for additional levels of gene regulation.

Translators of the histone code

Bromodomain proteins

Support for the histone code hypothesis came from discoveries that proteins with defined molecular functions recognize specific histone modifications, thereby translating the code into a biological process. Studies of the bromodomain, a conserved domain found in a variety of chromatin associated proteins, show that it specifically interacts with acetylated lysines of histone H3 and H4³⁸⁻⁴⁰. Bromodomain-containing proteins, such as GCN5, Brahma and TAF_{II}250, mediate several important cellular functions, such as transcriptional activation, maintenance of transcriptional activation through mitosis, and prevention of the spreading of silent chromatin^{39, 40}. For example, TAF_{II}250 interacts with the acetylated lysine 8 residue of histone H4 and the acetylated lysine 9 and 14 residues on histone H3³⁹. Association of TAF_{II}250 with acetylated histones results in the recruitment of TF_{II}D and activation of nearby gene expression⁴⁰.

Acetylated histone tails are not the only substrate for interactions; bromodomain-containing proteins also interact with other proteins that contain acetylated lysine

residues⁴¹. In many cases, these proteins are DNA-binding transcriptional activators. Interactions with bromodomain containing proteins are hypothesized to stabilize binding to gene regulatory regions, thereby promoting gene expression⁴¹.

Heterochromatin Protein 1 and Polycomb

Just as the bromodomain translates histone acetylation into gene activation, the chromo domain (CD) translates histone methylation into gene silencing. Chromo domain proteins such as Heterochromatin Protein 1 (HP1) and Polycomb (Pc) bind methylated lysine residues on the histone H3 tail. The genes encoding HP1 and Pc were connected to gene silencing through genetic screens carried out in *Drosophila*. A screen for modifiers of PEV identified the *Su(var)2-5* gene that encodes HP1⁴², and a screen for mutations that give rise to homeotic phenotypes led to the discovery of *Pc*⁴³. While HP1 and Pc Group (PcG) proteins, such as Polyhomeotic (Ph), are clearly involved in gene silencing, they localize to non-overlapping sites in the genome, as visualized by immunostaining on *Drosophila* polytene chromosomes (Figure 2), suggesting distinct targets.

The amino acid sequences of HP1 proteins are conserved among species. This conservation primarily resides within two domains, the amino-terminal CD and the structurally similar carboxy-terminal chromo shadow domain (CSD)⁴⁴. The CD and CSD are separated by a poorly conserved hinge region (Figure 3A). A hydrophobic pocket formed by the CD serves as an interaction site for di- and tri-methylated lysine nine residues on histone H3^{45, 46} (Figure 3C). This modification is generated by the SET-

domain histone methyltransferase SU(VAR)3-9. Association with the methyl mark is clearly the mechanism used for localization of the majority of HP1 to centric locations, however, recent data suggest that HP1 uses alternate mechanisms at other genomic sites⁴⁷⁻⁴⁹.

The hinge region is unstructured in solution and is thought to enable the chromo and chromo shadow domains to move independently of each other^{50,51}. Although the hinge region is unstructured, it plays a role in HP1 function as a nuclear localization signal required for active transport of the protein into the nucleus⁵². In addition, the hinge interacts with DNA and other nuclear proteins, such as histone H1 and HDACs, which function in gene silencing⁵³⁻⁵⁶. The hinge of HP1 has also been implicated in RNA binding, which appears to be required for pericentric heterochromatin localization in mammalian cells⁵⁷.

The CSD is structurally similar to the CD, also forming a hydrophobic pocket⁵⁸. In contrast to the CD, the CSD pocket has not been found to bind to modified histone tails, rather the CSD dimerizes with another CSD on a second molecule of HP1 through its α -helical region. Homodimerization generates a surface for interaction with a variety of nuclear proteins that contain a pentapeptide motif, PxVxL^{50,58}. These proteins possess a wide variety of functions, including chromatin remodeling, DNA replication, and transcription control^{50,59-61}. The role of HP1 in connection with many of these proteins is not well understood.

What is the mechanism of HP1-mediated gene silencing? Two approaches have been taken to address this question using *Drosophila*. In one approach a euchromatic transgene was mobilized into both euchromatic and heterochromatic regions of the genome. Transcription and chromatin structure analysis has been performed on the transgenes in both types of environments. Silenced transgenes inserted into heterochromatin appear less accessible to nuclease digestion and package into more regular nucleosome arrays than the expressed euchromatic transgenes^{62, 63}. Lower doses of HP1 relieved silencing of transgenes inserted into heterochromatin and produced a more open chromatin structure⁶⁴. Thus, one possible role of HP1 in heterochromatin is to compact chromatin, possibly through CSD-CSD interactions, preventing regulatory factors from associating with their target sites. Supporting this idea, high-resolution footprint analyses revealed a lack of transcription factors and paused polymerase at the promoters of transgenes inserted into heterochromatic regions, whereas association at the promoters of transgenes inserted into euchromatin was evident²⁵. Thus, HP1-mediated gene silencing appears to function at the level of transcription initiation prior to the recruitment of *trans*-acting factors (Figure 4A). Consistent with these findings, studies in mammals show that transcription factors associated with genes on the active X-chromosome are absent from genes on the Xi⁶⁵.

The second approach taken in *Drosophila* to address mechanisms of HP1-mediated silencing employs a tethering system. In this system, HP1 was fused to the DNA binding domain of the LacI repressor from *E. coli*^{66, 67}. The LacI-HP1 fusion protein was expressed in flies carrying *lac* operator repeats upstream of a reporter gene. Association

of LacI-HP1 with the *lac* operator repeats resulted in silencing of the reporter gene and corresponded to alterations in chromatin structure, including the generation of regular nucleosome arrays^{66,67}. These studies demonstrate that HP1 is sufficient to nucleate the formation of silent chromatin, even in the absence of repeat sequences typically found within heterochromatin.

Polycomb protein (Pc) associates with chromatin in a manner similar to that of HP1. The chromo domain of Pc forms a hydrophobic cleft that binds tri-methylated lysine 27 of histone H3 with high affinity (Figure 3D), and tri-methylated lysine 9 of histone H3 with lower affinity⁶⁸. Methylation at these residues is carried out by E(Z), a component of the Esc-E(Z) complex⁶⁹⁻⁷¹. Based on structural studies, it appears that Pc dimerizes and that two Pc molecules are unlikely to associate with one nucleosome (i.e. on both H3 tails within a nucleosome)⁶⁸. Therefore, two Pc molecules associated with adjacent nucleosomes might dimerize to compact chromatin, bringing about gene silencing⁶⁸.

What is the mechanism of Pc-mediated gene silencing? To address this question, a *Drosophila* transgene system was developed⁷². A reporter gene with an upstream Polycomb Response Element (PRE), the *cis*-element that targets the Pc complex, was inserted into the genome. Association of the Pc complex with the PRE led to gene silencing and a reduction in nuclease accessibility of the transgene promoter region⁷². Surprisingly, chromatin immunoprecipitation experiments revealed that *trans*-acting factors remained associated with regulatory elements in the transgene and that the paused polymerase maintained association with the silenced promoter. However, evidence of the

polymerase being engaged for transcription was absent⁷². Thus, Pc silencing is likely to block steps in initiation after recruitment of the polymerase (Figure 4B).

In summary, there are many similarities and differences between HP1- and Pc-mediated gene silencing. Both HP1 and Pc interact with chromosomes through a methylated residue of a histone H3 tail, though utilizing different residues. Biophysical and *in vivo* localization studies support the idea that both HP1 and Pc homodimerize, which might allow for inter-nucleosomal interactions^{58, 68}. HP1 and Pc both have a conserved carboxy domain, the CSD and Pc box, respectively, that interacts with a wide variety of nuclear factors^{48, 73} (Figure 3A, B). Thus, HP1 and Pc can be thought of as bridging molecules that link the chromosome with other factors required for proper chromosome dynamics. While HP1 and Pc use a similar mechanism for localization, cytological and genetic studies demonstrate that they have different targets. In addition, their mechanisms of silencing appear to be distinct. HP1-mediated silencing appears to prevent association of *trans*-acting factors, whereas Pc-mediated silencing allows association, but not activation (Figure 4). One explanation for this apparent difference in mechanism of silencing might be a reflection of when each protein acts in development. HP1 is presumed to act early in embryogenesis at the time when heterochromatin forms, and Pc acts later in embryogenesis to silence homeotic gene expression at subsequent developmental stages.

Connections between gene silencing and RNA

Non-coding RNAs

In recent years there is overwhelming support for a role of RNAs in the regulation of gene silencing. In this section we will focus on RNAs that carry out a role in gene silencing by association with the silenced gene itself. These RNAs are typically non-coding (ncRNAs), meaning that they lack protein-encoding capacity.

Non-coding RNAs play a pivotal role in X-chromosome inactivation. In undifferentiated cells *X-inactive specific transcript (Xist)* initiates from the X-inactivation center (Xic)¹¹. This 17 kb message is spliced, polyadenylated and spreads along the X-chromosome in *cis*. In addition to *Xist*, *Tsix* a second ncRNA, is transcribed from the opposite strand along the length of the *Xist* gene⁷⁴. At the onset of differentiation, *Xist* is exclusively expressed from the X-chromosome selected for inactivation, while *Tsix* is exclusively expressed from the X-chromosome that will remain active. Expression of *Xist* triggers a series of epigenetic modifications on the Xi that correlate with gene silencing (Figure 5A). One of the first events to occur is the *Xist*-dependent transient association of the Eed-Enx1 complex, a homologue of the Polycomb complex ESC-E(Z), that trimethylates lysine 27 of histone H3^{75,76}. Subsequent histone modifications include hypoacetylation of histone H3 and H4 and methylation of lysine 9 of histone H3⁷⁷. As predicted by the histone code, HP1 is enriched on the Xi¹³. A late event that accompanies stable expression of *Xist* is the enrichment of the non-replicative replacement histone variant macroH2A and DNA methylation. macroH2A and many of the histone

modifications specific for the Xi form patterns of domains along the length of Xi⁷⁷. Interestingly, domains on the Xi containing escaper genes, such as the pseudoautosomal region, show reduced amounts of the histone modifications associated with silencing¹⁴. Taken together, the combination of ncRNA, histone modifications and histone variants might be evidence for further complexity in the code, which is translated into domains of genes with different levels of expression.

In addition to their role in X-inactivation, non-coding RNAs are involved in genomic imprinting. The *Igf2r* gene cluster contains three protein-coding imprinted genes, including *Igf2r*, that are expressed when maternally inherited⁷⁸. Within this cluster, the ncRNA *Air* is only expressed when paternally inherited. *Air* expression correlates with silencing of all three protein-coding genes. Transcription of *Air* originates from an intron of *Igf2r* and is transcribed in the anti-sense orientation along the length of *Igf2r*, including the promoter region⁷⁸. Early studies of *Air* suggested promoter occlusion or anti-sense RNA mechanisms for silencing of *Igf2r*, with subsequent spreading of silent chromatin over the additional genes in the cluster⁷⁸. More recent studies rule out these possibilities and suggest a different, but related, mechanism. Initial silencing of *Igf2r* and transcriptional overlap between *Igf2r* and *Air* is not required for proper imprinting⁷⁹. New models of silencing within the *Igf2r* gene cluster liken the role of *Air* to *Xist*, suggesting spreading in *cis* within a domain, rather than along an entire chromosome⁷⁹. Association of *Air* with the *Igf2r* cluster is accompanied by chromatin modifications, such as deacetylation and histone methylation⁸⁰. However, differences between *Air* and *Xist*-mediated silencing have already been discovered. Mutations in *Eed* do not

affect silencing at the *Igf2r* locus⁸¹, suggesting that histone methylation by this Polycomb group complex is not linked with this imprinted domain as it is with X-chromosome inactivation.

RNAi-based initiation of silencing.

Evidence of RNAi was first studied in detail in *C. elegans*. Injection of anti-sense RNA corresponding to the *par-1* gene led to knock-down of *par-1* expression, as anticipated⁸². Unexpectedly, however, injection of the sense-strand control RNA had the same effect⁸². Furthermore, injection of double stranded RNA (dsRNA) for a given gene was more effective in silencing the gene than injection of either strand alone⁸³. Components of the pathway involved in this phenomenon, called RNA interference (RNAi), have recently been elucidated. dsRNA is recognized by the Dicer enzyme that cleaves the RNA into small interfering RNA (siRNA) molecules of approximately 22 nucleotides in length⁸⁴. The siRNA associates with the RISC complex and is unwound into single stranded antisense RNA, activating the RISC complex. Activated RISC associates with the target mRNA, with the single stranded antisense RNA annealing to the region of homology, causing cleavage and degradation⁸⁴. The outcome of this process, referred to as RNAi results in gene silencing⁸⁴.

The discovery of the RNAi mediated gene-silencing pathway led to the identification of natural gene targets. In *Drosophila*, tandem repeated *Stellate* genes, encoding a regulatory subunit for casein kinase II, reside in a heterochromatic cluster on the X-chromosome⁸⁵. The *Stellate* repeats are silenced early in *Drosophila* development by

sense and anti-sense transcripts produced from homologous sequences at the *Suppressor of Stellate* [*Su(ste)*] locus on the Y chromosome⁸⁶. Lack of silencing of the *Stellate* repeats in male testes leads to meiotic abnormalities and male sterility⁸⁵. Mutations in *aubergine* (*aub*) and *homeless* (*hls*), both components of the RISC complex, relieve silencing of the *Stellate* repeats^{85, 86}, demonstrating a role for RNAi in heterochromatic gene silencing.

Likewise, repetitive sequences located near *S. pombe* centromeres are natural targets of the RNAi pathway⁸⁷. The formation of silenced chromatin encompassing these repeats occurs in a two-step process involving initiation and spreading (Figure 5B). Initiation requires processing of the dsRNAs, from the bi-directionally transcribed centric repeats, through the RNAi pathway. These siRNAs, in conjunction with RISC, associate with their homologous sequences within centric regions. Initiation of silencing occurs through the recruitment of histone modifying proteins, resulting in the methylation of lysine 9 of histone H3 that serves as a docking site for Swi6, the *S. pombe* HP1 ortholog⁸⁷.

Consistent with this process, mutations in the RNAi pathway lead to the loss of Swi6 at the centric repeats⁸⁷. Spreading of silent chromatin is independent of the RNAi pathway and is hypothesized to involve the histone code-generating machinery. A similar relationship exists between RNAi and centric heterochromatin in *Drosophila*⁸⁸. In contrast to the post-transcriptional gene silencing of *Stellate* repeats described above, silencing near centromeres appears to occur at the transcriptional level, through the formation of silent chromatin initiated by siRNAs.

Conclusions

Now that the genomes of many organisms have been sequenced, it is our challenge to discover the rules that regulate gene expression. The histone code appears to be a good candidate for predicting patterns of gene expression. The code continues to grow in complexity with both the identification of new histone modifications and knowledge about how these modifications influence each other and the binding of non-histone chromatin proteins^{89,90}. The histone code might be initially established at locations within the genome by the distribution of DNA sequence elements. These elements attract DNA-binding proteins that further recruit chromatin-modifying machinery, thereby setting up the histone code. Other factors that might influence gene expression on a domain level include the arrangement of chromatin within the three-dimensional nucleus⁹¹. In some cases, positioning of chromosomes within the nucleus and localization of specific RNA, such as *Xist*, are required for gene silencing. Recent data clearly demonstrate a role for RNA in the initiation of silent chromatin. Beyond initiation, a common property of silent chromatin is the ability to “spread” from the initiation site. Mechanisms responsible for this process are not well defined, and are the focus of future studies.

Acknowledgements

We would like to thank members of the Wallrath lab for a critical reading of the manuscript. We would also like to thank Greta Schrift and Yuhong Li for assistance with figures. Research on HP1 in *Drosophila* and humans is supported by grants NIH

GM61513 and Department of Defense Breast Cancer Research grant DAMD17-02-1-0424), respectively.

References

1. Stargell, L. A. and Struhl, K. *Trends Genet*, **1996**, 12, 311.
2. Muller, H. J. *Journal of Genetics*, **1930**, 22, 299.
3. Weiler, K. S. and Wakimoto, B. T. *Annu Rev Genet*, **1995**, 29, 577.
4. Grigliatti, T. *Methods Cell Biol*, **1991**, 35, 587.
5. Reuter, G. and Spierer, P. *Bioessays*, **1992**, 14, 605.
6. Allshire, R. C., Javerzat, J. P., Redhead, N. J. and Cranston, G. *Cell*, **1994**, 76, 157.
7. Ayoub, N., Goldshmidt, I. and Cohen, A. *Genetics*, **1999**, 152, 495.
8. Gottschling, D. E., Aparicio, O. M., Billington, B. L. and Zakian, V. A. *Cell*, **1990**, 63, 751.
9. Festenstein, R., Sharghi-Namini, S., Fox, M., Roderick, K., Tolaini, M., Norton, T., Saveliev, A., Kioussis, D. and Singh, P. *Nat Genet*, **1999**, 23, 457.
10. Saveliev, A., Everett, C., Sharpe, T., Webster, Z. and Festenstein, R. *Nature*, **2003**, 422, 909.
11. Park, Y. and Kuroda, M. I. *Science*, **2001**, 293, 1083.
12. Barr, M. L. *Int Rev Cytol*, **1966**, 19, 35.
13. Chadwick, B. P. and Willard, H. F. *Hum Mol Genet*, **2003**, 12, 2167.
14. Brown, C. J. and Greally, J. M. *Trends Genet*, **2003**, 19, 432.
15. Brown, C. J. and Robinson, W. P. *Clin Genet*, **2000**, 58, 353.
16. Burn, J., Povey, S., Boyd, Y., Munro, E. A., West, L., Harper, K. and Thomas, D. *J Med Genet*, **1986**, 23, 494.
17. Scott, R. J. and Spielman, M. *Curr Biol*, **2004**, 14, R201.

18. Walter, J. and Paulsen, M. *Hum Mol Genet*, **2003**, 12 Spec No 2, R215.
19. Walter, J. and Paulsen, M. *Semin Cell Dev Biol*, **2003**, 14, 101.
20. Verona, R. I., Mann, M. R. and Bartolomei, M. S. *Annu Rev Cell Dev Biol*, **2003**, 19, 237.
21. Khorasanizadeh, S. *Cell*, **2004**, 116, 259.
22. Wolffe, A. P., Khochbin, S. and Dimitrov, S. *Bioessays*, **1997**, 19, 249.
23. Grigoryev, S. A. *FEBS Lett*, **2004**, 564, 4.
24. Hill, R. J., Watt, F., Wilson, C. M., Fifis, T., Underwood, P. A., Tribbick, G., Geysen, H. M. and Thomas, J. O. *Chromosoma*, **1989**, 98, 411.
25. Cryderman, D. E., Tang, H., Bell, C., Gilmour, D. S. and Wallrath, L. L. *Nucleic Acids Res*, **1999**, 27, 3364.
26. Edmondson, D. G. and Roth, S. Y. *FASEB J*, **1996**, 10, 1173.
27. Li, Q., Wrangé, O. and Eriksson, P. *Int J Biochem Cell Biol*, **1997**, 29, 731.
28. Wallrath, L. L. *Curr Opin Genet Dev*, **1998**, 8, 147.
29. Wang, Y., Fischle, W., Cheung, W., Jacobs, S., Khorasanizadeh, S. and Allis, C. D. *Novartis Found Symp*, **2004**, 259, 3.
30. Osley, M. A. *Biochim Biophys Acta*, **2004**, 1677, 74.
31. Jenuwein, T. and Allis, C. D. *Science*, **2001**, 293, 1074.
32. Kurdistani, S. K. and Grunstein, M. *Nat Rev Mol Cell Biol*, **2003**, 4, 276.
33. Wade, P. A. *Bioessays*, **2001**, 23, 1131.
34. Freitag, M., Hickey, P. C., Khlaifallah, T. K., Read, N. D. and Selker, E. U. *Mol Cell*, **2004**, 13, 427.

35. Soppe, W. J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M. S., Jacobsen, S. E., Schubert, I. and Fransz, P. F. *EMBO J*, **2002**, 21, 6549.
36. Cervoni, N. and Szyf, M. *J Biol Chem*, **2001**, 276, 40778.
37. Ehrenhofer-Murray, A. E. *Eur J Biochem*, **2004**, 271, 2335.
38. Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K. and Zhou, M. M. *Nature*, **1999**, 399, 491.
39. Bottomley, M. J. *EMBO Rep*, **2004**, 5, 464.
40. Marmorstein, R. and Berger, S. L. *Gene*, **2001**, 272, 1.
41. Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sachchidanand, Sanchez, R., Zeleznik-Le, N. J., Ronai, Z. and Zhou, M. M. *Mol Cell*, **2004**, 13, 251.
42. Eissenberg, J. C., James, T. C., Foster-Hartnett, D. M., Hartnett, T., Ngan, V. and Elgin, S. C. *Proc Natl Acad Sci U S A*, **1990**, 87, 9923.
43. Lewis, E. B. *Nature*, **1978**, 276, 565.
44. Eissenberg, J. C. and Elgin, S. C. *Curr Opin Genet Dev*, **2000**, 10, 204.
45. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. *Nature*, **2001**, 410, 116.
46. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. and Kouzarides, T. *Nature*, **2001**, 410, 120.
47. Greil, F., van der Kraan, I., Delrow, J., Smothers, J. F., de Wit, E., Bussemaker, H. J., van Driel, R., Henikoff, S. and van Steensel, B. *Genes Dev*, **2003**, 17, 2825.
48. Li, Y., Kirschmann, D. A. and Wallrath, L. L. *Proc Natl Acad Sci U S A*, **2002**, 99 Suppl 4, 16462.

49. Perrini, B., Piacentini, L., Fanti, L., Altieri, F., Chichiarelli, S., Berloco, M., Turano, C., Ferraro, A. and Pimpinelli, S. *Mol Cell*, **2004**, 15, 467.
50. Thiru, A., Nietlispach, D., Mott, H. R., Okuwaki, M., Lyon, D., Nielsen, P. R., Hirshberg, M., Verreault, A., Murzina, N. V. and Laue, E. D. *EMBO J*, **2004**, 23, 489.
51. Brasher, S. V., Smith, B. O., Fogh, R. H., Nietlispach, D., Thiru, A., Nielsen, P. R., Broadhurst, R. W., Ball, L. J., Murzina, N. V. and Laue, E. D. *EMBO J*, **2000**, 19, 1587.
52. Smothers, J. F. and Henikoff, S. *Mol Cell Biol*, **2001**, 21, 2555.
53. Badugu, R., Shareef, M. M. and Kellum, R. *J Biol Chem*, **2003**, 278, 34491.
54. Nielsen, A. L., Oulad-Abdelghani, M., Ortiz, J. A., Remboutsika, E., Chambon, P. and Losson, R. *Mol Cell*, **2001**, 7, 729.
55. Zhang, C. L., McKinsey, T. A. and Olson, E. N. *Mol Cell Biol*, **2002**, 22, 7302.
56. Meehan, R. R., Kao, C. F. and Pennings, S. *EMBO J*, **2003**, 22, 3164.
57. Muchardt, C., Guilleme, M., Seeler, J. S., Trouche, D., Dejean, A. and Yaniv, M. *EMBO Rep*, **2002**, 3, 975.
58. Cowieson, N. P., Partridge, J. F., Allshire, R. C. and McLaughlin, P. J. *Curr Biol*, **2000**, 10, 517.
59. Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K. and Rauscher, F. J., 3rd. *Genes Dev*, **2003**, 17, 1855.
60. Murzina, N., Verreault, A., Laue, E. and Stillman, B. *Mol Cell*, **1999**, 4, 529.
61. Ye, Q. and Worman, H. J. *J Biol Chem*, **1996**, 271, 14653.
62. Wallrath, L. L. and Elgin, S. C. *Genes Dev*, **1995**, 9, 1263.
63. Sun, F. L., Cuaycong, M. H. and Elgin, S. C. *Mol Cell Biol*, **2001**, 21, 2867.

64. Cryderman, D. E., Cuaycong, M. H., Elgin, S. C. and Wallrath, L. L. *Chromosoma*, **1998**, 107, 277.
65. Komura, J., Sheardown, S. A., Brockdorff, N., Singer-Sam, J. and Riggs, A. D. *J Biol Chem*, **1997**, 272, 10975.
66. Li, Y., Danzer, J. R., Alvarez, P., Belmont, A. S. and Wallrath, L. L. *Development*, **2003**, 130, 1817.
67. Danzer, J. R. and Wallrath, L. L. *Development*, **2004**, 131, 3571.
68. Min, J., Zhang, Y. and Xu, R. M. *Genes Dev*, **2003**, 17, 1823.
69. Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E. and Simon, J. A. *Cell*, **2002**, 111, 197.
70. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. and Zhang, Y. *Science*, **2002**, 298, 1039.
71. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. *Genes Dev*, **2002**, 16, 2893.
72. Dellino, G. I., Schwartz, Y. B., Farkas, G., McCabe, D., Elgin, S. C. and Pirrotta, V. *Mol Cell*, **2004**, 13, 887.
73. Jones, D. O., Cowell, I. G. and Singh, P. B. *Bioessays*, **2000**, 22, 124.
74. Lee, J. T., Davidow, L. S. and Warshawsky, D. *Nat Genet*, **1999**, 21, 400.
75. Silva, J., Mak, W., Zvetkova, I., Appanah, R., Nesterova, T. B., Webster, Z., Peters, A. H., Jenuwein, T., Otte, A. P. and Brockdorff, N. *Dev Cell*, **2003**, 4, 481.
76. Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., Otte, A. P., Panning, B. and Zhang, Y. *Science*, **2003**, 300, 131.
77. Hall, L. L. and Lawrence, J. B. *Semin Cell Dev Biol*, **2003**, 14, 369.

78. Sleutels, F., Zwart, R. and Barlow, D. P. *Nature*, **2002**, 415, 810.
79. Sleutels, F., Tjon, G., Ludwig, T. and Barlow, D. P. *EMBO J*, **2003**, 22, 3696.
80. Yang, Y., Li, T., Vu, T. H., Ulaner, G. A., Hu, J. F. and Hoffman, A. R. *Endocrinology*, **2003**, 144, 5658.
81. Mager, J., Montgomery, N. D., de Villena, F. P. and Magnuson, T. *Nat Genet*, **2003**, 33, 502.
82. Guo, S. and Kempthues, K. J. *Cell*, **1995**, 81, 611.
83. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. *Nature*, **1998**, 391, 806.
84. Denli, A. M. and Hannon, G. J. *Trends Biochem Sci*, **2003**, 28, 196.
85. Gvozdev, V. A., Aravin, A. A., Abramov, Y. A., Klenov, M. S., Kogan, G. L., Lavrov, S. A., Naumova, N. M., Olenkina, O. M., Tulin, A. V. and Vagin, V. V. *Genetica*, **2003**, 117, 239.
86. Aravin, A. A., Klenov, M. S., Vagin, V. V., Bantignies, F., Cavalli, G. and Gvozdev, V. A. *Mol Cell Biol*, **2004**, 24, 6742.
87. Grewal, S. I. and Rice, J. C. *Curr Opin Cell Biol*, **2004**, 16, 230.
88. Pal-Bhadra, M., Leibovitch, B. A., Gandhi, S. G., Rao, M., Bhadra, U., Birchler, J. A. and Elgin, S. C. *Science*, **2004**, 303, 669.
89. Freitas, M. A., Sklenar, A. R. and Parthun, M. R. *J Cell Biochem*, **2004**, 92, 691.
90. Syka, J. E., Marto, J. A., Bai, D. L., Horning, S., Senko, M. W., Schwartz, J. C., Ueberheide, B., Garcia, B., Busby, S., Muratore, T., Shabanowitz, J. and Hunt, D. F. *J Proteome Res*, **2004**, 3, 621.
91. Pederson, T. *Curr Opin Genet Dev*, **2004**, 14, 203.

Figure Legends

Figure 1. Position effect variegation across species. A. In *Drosophila*, PEV is observed when a chromosomal translocation moves the *white* gene near a centromere. This results in a white and red variegated eye phenotype. *Su(var)* mutations result in increased expression of the *white* gene. *E(var)* mutations result in decreased expression of the *white* gene. B. In *S. pombe*, PEV is observed when the *ade6* transgene is placed near a centromere. *ade6* silencing leads to red pigmentation, therefore, PEV results in red and white sectorial colonies. C. In mouse cells, PEV is observed when a *CD2* transgene is inserted near a centromere. Variegation results in a subpopulation of the cells lacking the CD2 transmembrane receptor protein.

Figure 2. HP1 and Pc proteins do not co-localize on *Drosophila* polytene chromosomes. Polytene chromosomes from *Drosophila* third instar larval salivary glands that were squashed and stained with antibodies to HP1 (red) and Ph, a PcG protein (green). The chromocenter, the site of fusion of all centromeres, is labeled C.

Figure 3. Structure of HP1 and Pc. A. Diagram of the domain structure of HP1 protein, showing the chromo domain, the hinge region, and the chromoshadow domain. B. Diagram of the domain structure of Pc, showing the chromo domain and the Pc box. C. The chromo domain (1KNA.pdb) of HP1 (blue) forms a hydrophobic pocket around the di-methylated lysine nine (green) of the histone H3 peptide (red). D. The chromo

domain (1PDQ.pdb) of Pc (blue) forms a cleft that specifically interacts with trimethylated lysine 27 (green) of the histone H3 peptide (red).

Figure 4. Mechanisms of HP1 and Pc gene silencing. A. One model for HP1-mediated gene silencing involves HP1 binding to methylated histones and forming a regular array of nucleosomes that blocks transcriptional machinery from gaining access to the gene promoter. B. One model for Pc-mediated gene silencing involves the Pc-complex interacting with transcription factors at the gene promoter and preventing engagement of the polymerase.

Figure 5. Examples of ncRNA and RNAi-mediated gene silencing. A. An example of ncRNA-based silencing is X-inactivation. *Xist* RNA (pink line) is expressed from the Xi, triggering a series of histone and DNA modifications that accompany Eed-Enx1 and HP1 binding. B. In *S. pombe*, siRNAs corresponding to repetitive sequences located near centromeres, in association with the RISC complex, trigger initiation of heterochromatin. Ctr4 methylates lysine 9 of histone H3 and attracts the Swi6/HP1 protein.

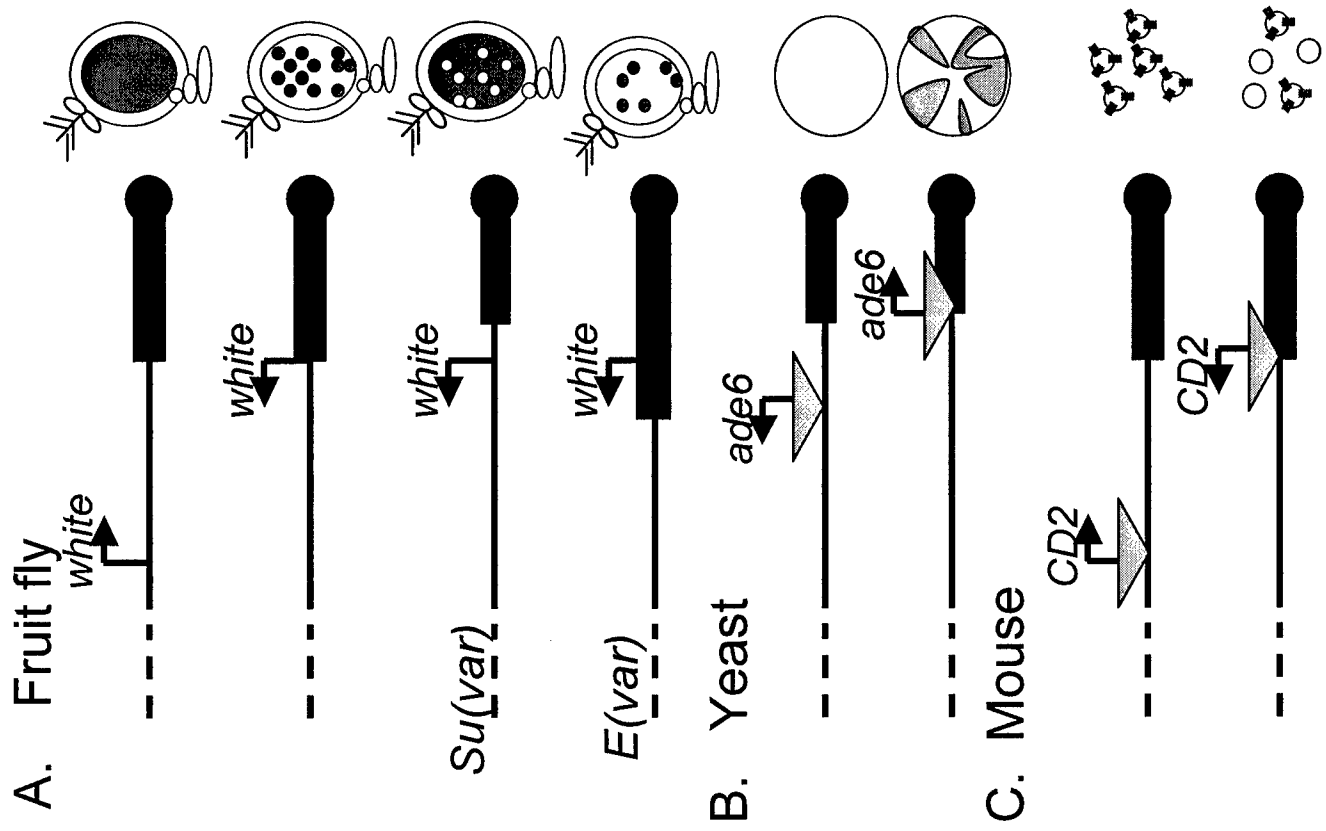
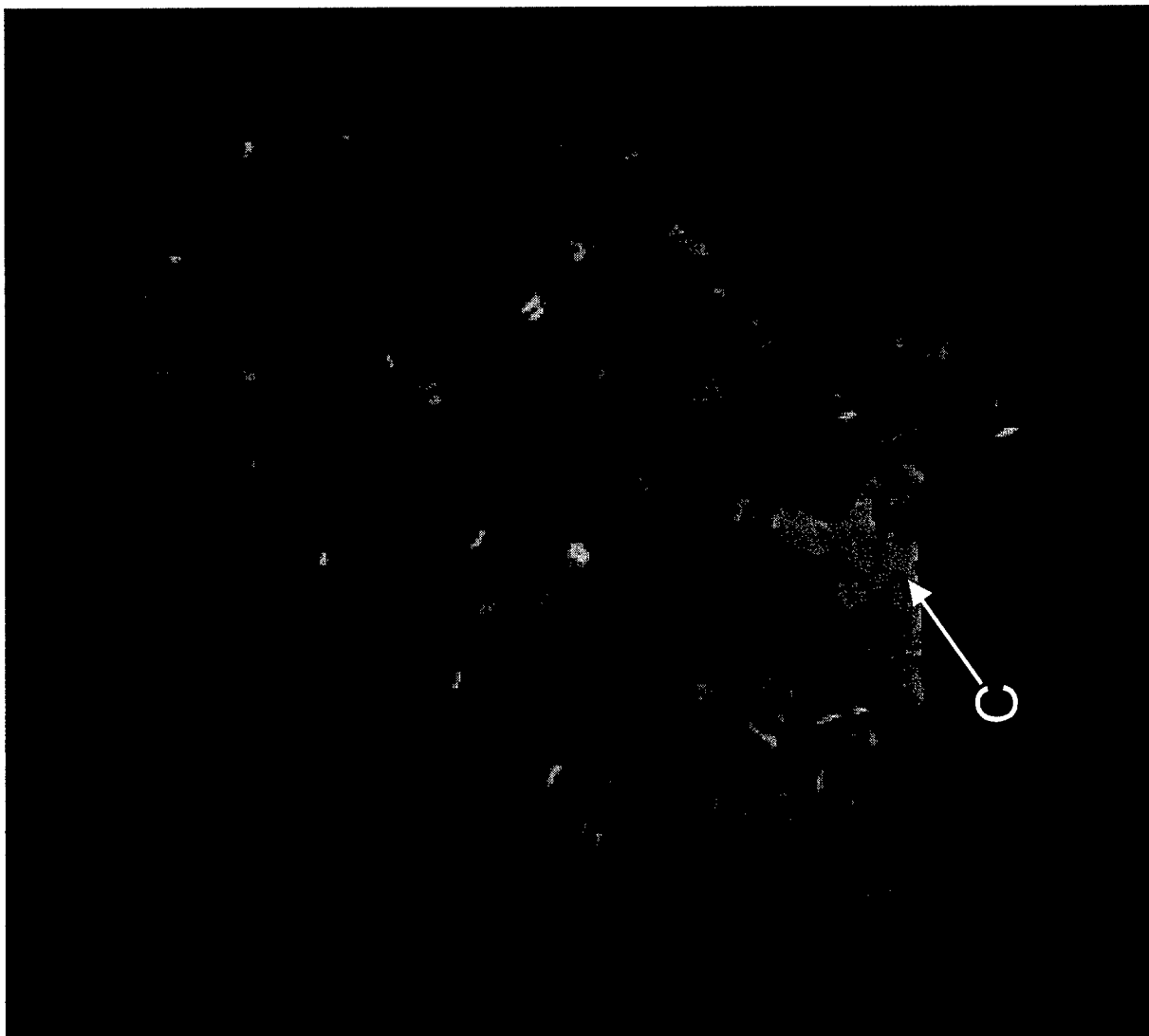


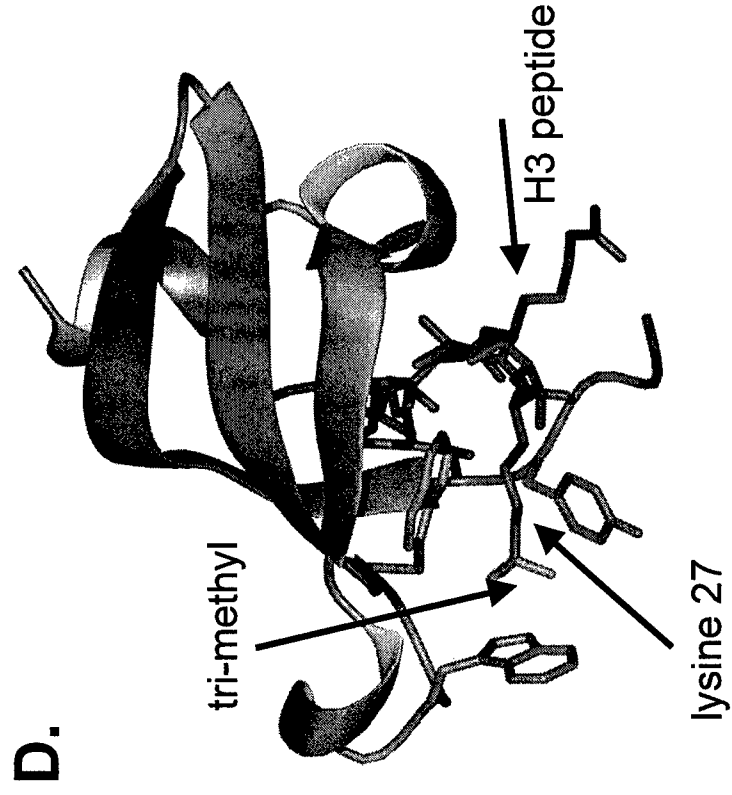
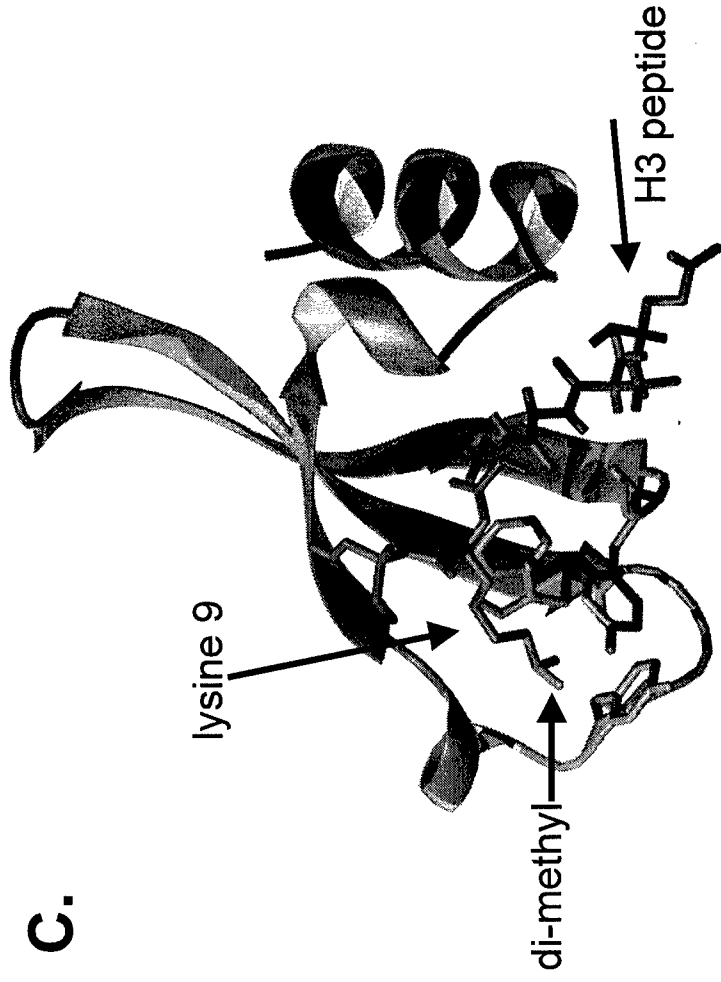
Table 1. Examples of epigenetic modifications and their consequences on nuclear events

Substrate	Modification	Associated function
DNA	Methylation of CpG	Transcriptional repression
histone H3	Methylated K4	Transcriptional activation
	Acetylated K9	Transcriptional activation
	Methylated K9	Transcriptional repression by HP1
	Phosphorylated S10	Chromosome condensation and mitosis/meiosis
	Acetylated K14	Transcriptional activation
	Methylated K27	Transcriptional repression by Pc
	Methylated K79	Transcriptional activation, telomeric silencing
	Methylated K122	Unknown function
histone H4	Phosphorylated S1	Chromatin condensation during mitosis
	Methylated R3	Enriched at developmentally regulated genes
	Acetylated K16	Transcriptional activation
histone H2A	Phosphorylated S1	Transcriptional repression, chromatin condensation
	Phosphorylated T119	Cell cycle progression?
histone H2B	Phosphorylated S14	Apoptotic chromosome condensation
	Phosphorylated S33	Transcriptional activation
	Ubiquitylated K119 or K123	Histone H3 methylation
histone H1	Phosphorylation	Transcriptional activation

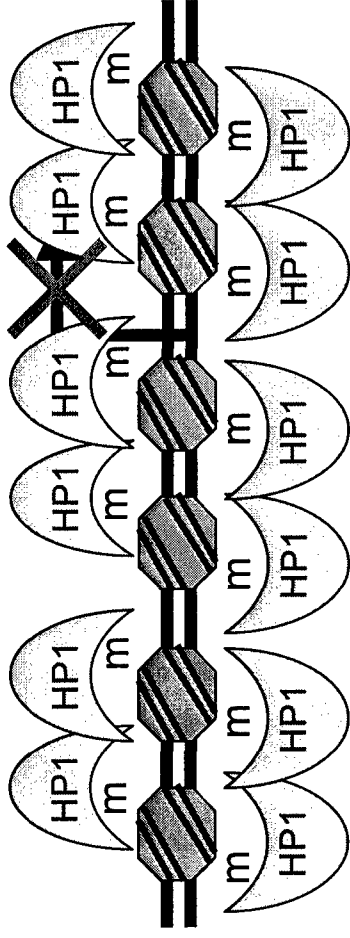


A. ☐ chromo ☐ hinge ☐ chromo shadow ☐

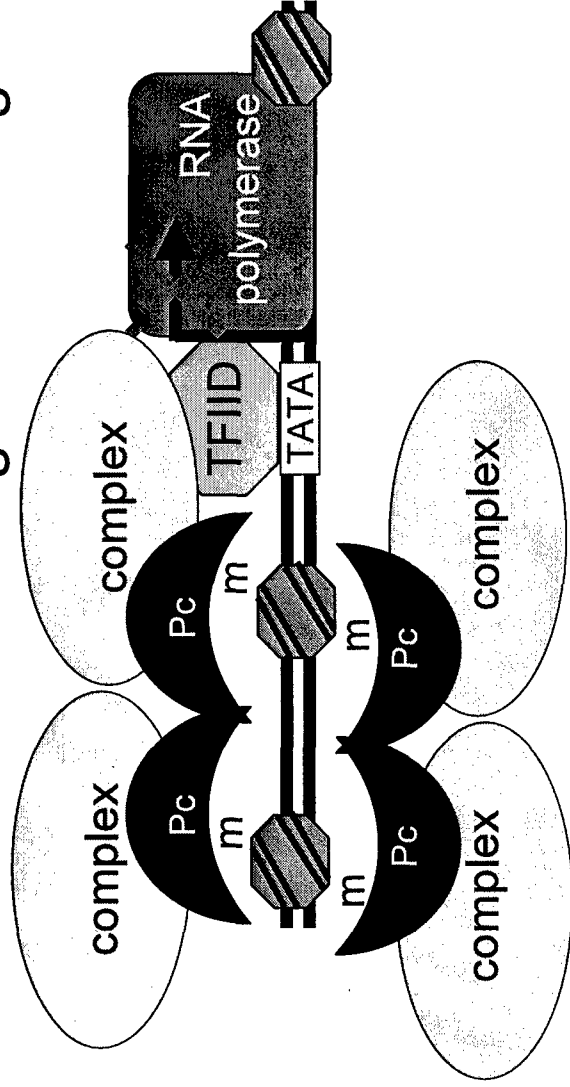
B. ☐ chromo ☐ ☐ Pc box ☐



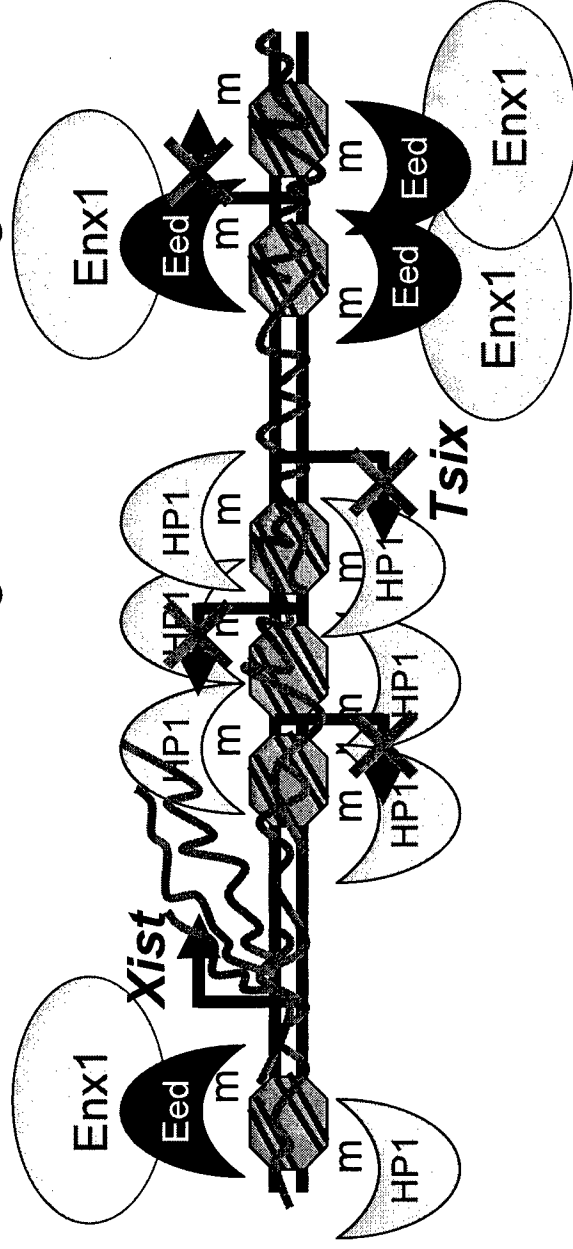
A. HP1-mediated gene silencing



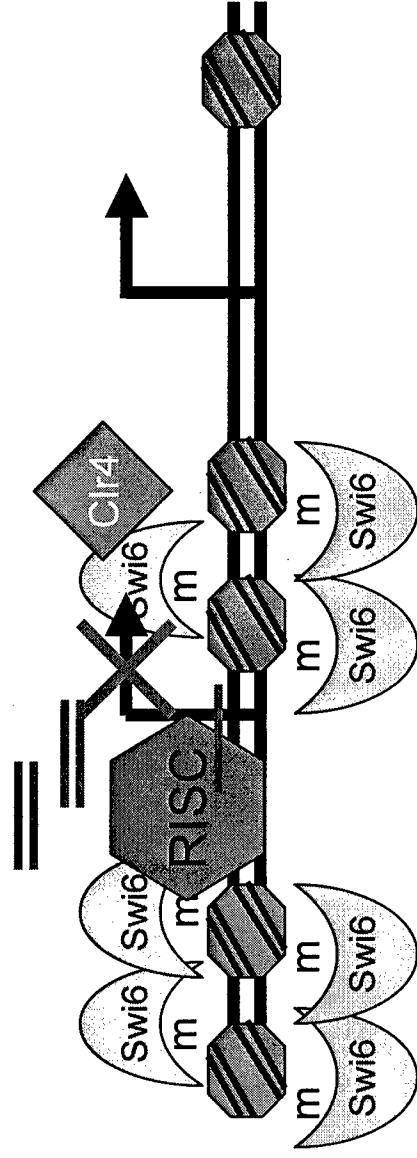
B. Pc-mediated gene silencing



A. Xist-mediated gene silencing



B. RNAi-based gene silencing



University of Iowa Genetics Program Retreat
University of Iowa IMU
September, 2004

Effects of knocking-down Heterochromatin Protein 1^{Hs} in breast cancer cells

Timothy J. Moss^{1,2,3}, Sara L. Sloat¹,

Mary J.C. Hendrix⁴, Dawn A. Kirschmann⁴ and Lori L. Wallrath^{1,3}

Department of Biochemistry¹, Medical Scientist Training Program²,

Genetics Program³, Department MED-Pediatrics Northwestern University⁴

Heterochromatin protein 1 (HP1) is a conserved chromosomal protein enriched in heterochromatic regions of the genome. In humans there are three HP1 family members, HP1^{Hs α} , HP1^{Hs β} and HP1^{Hs γ} . HP1 proteins contain two conserved domains, an amino chromo domain that associates with methylated lysine 9 of histone H3, and a carboxy chromo shadow domain that interacts with a variety of nuclear proteins. These domains are thought to regulate protein-protein interactions that influence gene expression. HP1^{Hs α} is down-regulated in invasive/metastatic breast cancer cells compared with poorly invasive/non-metastatic breast cancer cells. Our working hypothesis is that HP1 plays a role in invasion/metastasis through the regulation of gene expression. To address the *in vivo* function of HP1^{Hs α} in human breast cancer (MCF-7) cells, we have used two methods for RNAi knock-down, transfection of a pool of short, synthetic double-stranded RNA molecules, and infection with an adenovirus containing short hairpin RNAi sequences corresponding to HP1^{Hs α} . Independently, both methods have resulted in up to 99% HP1^{Hs α} protein knock-down. Preliminary experiments indicate the knock-down of HP1^{Hs α} is associated with changes in cellular morphology and adhesion. These changes include a decrease in cell clustering and a reduction in adherence. Initial studies indicate differences in the invasive potential of the knock-down MCF-7 cells based on *in vitro* invasion assays, implicating changes in adhesion and motility. To investigate the correlation between these phenotypes and the effects of HP1^{Hs α} on gene expression, a microarray analysis will be performed comparing cells with and without HP1^{Hs α} knock-down. Collectively, these studies will elucidate changes that occur in chromatin and gene expression during the transition in breast cancer associated with invasion and metastasis.

University of Iowa Medical Scientist Training Program Retreat
Backbone State Park, Iowa
August, 2004

The role of Heterochromatin Protein 1 in breast cancer cells

Timothy J. Moss^{1,2,3}, Sara L. Sloat¹,

Mary J.C. Hendrix⁴, Dawn A. Kirschmann⁴ and Lori L. Wallrath^{1,3}

Department of Biochemistry¹, Medical Scientist Training Program²,

Genetics Program³, Department MED-Pediatrics Northwestern University⁴

Heterochromatin protein 1 (HP1) is a conserved chromosomal protein enriched in heterochromatic regions of the genome, but also localizes to sites within euchromatin. In humans there are three HP1 family members, HP1^{Hsα}, HP1^{Hsβ} and HP1^{Hsγ}. HP1 proteins contain two conserved domains, an amino chromo domain that associates with methylated lysine 9 of histone H3, and a carboxy chromo shadow domain that interacts with a variety of nuclear proteins. These domains are thought to regulate protein-protein interactions that influence gene expression. HP1^{Hsα} is down-regulated in invasive/metastatic breast cancer cells compared with poorly invasive/non-metastatic breast cancer cells. Our working hypothesis is that HP1 plays a role in invasion/metastasis through the regulation of gene expression. To address the *in vivo* function of HP1^{Hsα} in human breast cancer (MCF-7) cells, we have used two methods for RNAi knock-down, transfection of a pool of short, synthetic double-stranded RNA molecules, and infection with an adenovirus containing short hairpin RNAi sequences corresponding to HP1^{Hsα}. Independently, both methods have resulted in up to 99% HP1^{Hsα} protein knock-down. Preliminary experiments indicate the knock-down of HP1^{Hsα} is associated with changes in cellular morphology and adhesion. These changes include a decrease in cell clustering and a reduction in adherence. Initial studies indicate differences in the invasive potential of the knock-down MCF-7 cells based on *in vitro* invasion assays, implicating changes in adhesion and motility. To investigate the correlation between these phenotypes and the effects of HP1^{Hsα} on gene expression, a microarray analysis will be performed comparing cells with and without HP1^{Hsα} knock-down. Collectively, these studies will elucidate changes that occur in chromatin and gene expression during the transition in breast cancer associated with invasion and metastasis.

Genetics Ph.D. Interdisciplinary Program Retreat, Iowa City, IA, September, 2004

Heterochromatin Protein 1: Development of a Novel Breast Cancer Metastasis Marker

Laura E. Norwood¹, Lindsay Wright¹, Naira Margaryan², Mary J.C. Hendrix², Dawn A. Kirschmann², **Lori L. Wallrath¹**

¹ Department of Biochemistry, University of Iowa

² Department of Anatomy and Cell Biology, University of Iowa

Metastasis is the process by which cancerous cells break away from a primary tumor and establish a tumor at a distant location within the body. The five-year survival rate for women with a primary breast cancer tumor is 97%. In contrast, the five-year survival rate for women with metastatic breast cancer is only 23%. Currently, there are few molecular markers for metastasis. We have determined that Heterochromatin protein 1^{Hsa} (HP1^{Hsa}) is down-regulated in highly invasive/metastatic breast cancer compared to poorly invasive/non-metastatic breast cancer. HP1 proteins are enriched at heterochromatic regions of the genome and play a role in gene silencing. HP1 proteins contain a chromo domain that interacts with methylated lysine nine of histone H3 and a chromoshadow domain that homodimerizes and interacts with a variety of nuclear proteins. We propose that in non-metastatic breast cancer cells, HP1^{Hsa} functions to silence genes that promote metastasis. To study the role of HP1^{Hsa} in breast cancer cells, we introduced exogenous *HP1^{Hsa}* by adenoviral infection into the highly invasive/metastatic MDA-MB-231 breast cancer cells. Expression of wild type *HP1^{Hsd}* in MDA-MB-231 cells reduces the invasive potential by 25%. Expression of either a nuclear-localized *EGFP* or a mutant *HP1^{Hsa}* (I165E) that fails to homodimerize does not change the invasive potential of the MDA-MB-231 cells. In contrast, introduction of a mutant *HP1^{Hsa}* (W174A) that fails to interact with known protein partners containing a PxVxL motif leads to a 25% decrease in invasion. Taken together, these data demonstrate that dimerization, but not interactions with proteins that contain a PxVxL motif, is required for HP1^{Hsa} to suppress invasion in metastatic breast cancer cells. Future studies will involve identification of HP1^{Hsa}-regulated genes required for the suppression of invasion.

AACR 96th Annual Meeting, Anaheim, CA,
April 2005

Heterochromatin Protein 1 (HP1^{Hsα}) as a candidate breast cancer metastasis suppressor.

Laura E. Norwood¹, Timothy J. Moss¹, Naira Margaryan², Sara Sloat¹, Lindsay Wright¹, Mary J.C. Hendrix², Dawn A. Kirschmann², and Lori L. Wallrath¹

¹ Department of Biochemistry, University of Iowa, Iowa City, Iowa

² Children's Memorial Research Center, Chicago, Illinois

An estimated 200,000 American women are diagnosed with invasive breast cancer each year. The five-year survival rate for women with primary breast cancer is 97%, but the rate is only 23% for women diagnosed with metastatic breast cancer. Currently, there is a need for molecular markers for metastasis. We have determined that the chromatin protein, HP1^{Hsα}, is down-regulated in highly invasive/metastatic breast cancer compared to poorly invasive/non-metastatic breast cancer. HP1^{Hsα} is one of three Heterochromatin Protein 1 (HP1) family members in mammals. HP1 proteins are structural components of centric heterochromatin and also regulate euchromatic gene expression through epigenetic mechanisms. HP1 family members have two domains, the chromo domain (CD) and the chromo shadow domain (CSD), separated by a flexible hinge region. The CD binds to methylated lysine nine of histone H3, an epigenetic mark for gene silencing. The CSD homodimerizes and generates a platform for interaction with a variety of nuclear proteins containing a PxVxL motif. We hypothesize that HP1^{Hsα} functions as a breast cancer metastasis suppressor through silencing genes that promote invasion and metastasis. We are using the poorly invasive/non-metastatic cell line, MCF-7, and the highly invasive/metastatic breast cancer cell line, MDA-MB-231, as a model system to study the role of HP1^{Hsα} in invasion and metastasis. Introduction of RNAi constructs causing knock-down of *HP1^{Hsα}* in MCF-7 cells increases their *in vitro* invasion by 50%. Consistent with this finding, introduction of exogenous *HP1^{Hsα}* into MDA-MB-231 cells reduces *in vitro* invasion by 29%. An amino acid substitution in the CSD (I165E) that disrupts HP1^{Hsα} homodimerization does not suppress invasion of MDA-MB-231 cells, but an amino acid substitution in the CSD (W174A) that disrupts the platform needed for interaction with proteins containing the PxVxL motif suppresses *in vitro* invasion by 32%. These results indicate that HP1^{Hsα} dimerization is required for suppression of invasion. The decrease in *in vitro* invasion caused by introduction of either wild-type HP1^{Hsα} or the W174A mutant correlates with an increase in levels of integrin $\alpha 3$ mRNA. Future studies include the identification of additional genes misregulated by altered HP1^{Hsα} dosage through candidate gene and microarray analyses. Understanding the regulation and function of HP1^{Hsα} in breast cancer invasion and metastasis will increase our knowledge of metastatic progression and may lead to new ways to diagnose and treat metastasis.

Short title: HP1 as a metastasis suppressor

University of Iowa, College of Medicine Research Week

May 2005

The effects of HP1Hsa knock-down on breast cancer invasion/metastasis. An estimated 200,000 American women are diagnosed with invasive breast cancer each year. The five-year survival rate for women with primary breast cancer is 98%, but the rate is only 26% for women diagnosed with metastatic breast cancer. Currently, the molecular mechanism of metastatic progression is not well understood. We have determined that heterochromatin protein 1, HP1Hsa, is down-regulated in highly invasive/metastatic breast cancer compared to poorly invasive/non-metastatic breast cancer. HP1 proteins are structural components of centric heterochromatin and also regulate gene expression through epigenetic mechanisms. HP1 family members have two domains, the chromo domain (CD) and the chromo shadow domain (CSD), separated by a flexible hinge region. The CD binds to methylated lysine nine of histone H3, an epigenetic mark for gene silencing. The CSD homodimerizes and generates a platform for interaction with a variety of nuclear proteins containing a PxVxL motif. We hypothesize that HP1-Hsalpha functions as a breast cancer metastasis suppressor by regulating genes involved in metastasis. A poorly invasive/non-metastatic cell line, MCF-7, was used as models to study the role of HP1Hsa in invasion and metastasis. Introduction of RNAi constructs causing knock-down of HP1Hsa in MCF-7 cells increases their in vitro invasion by 50%. Current studies include the identification of genes misregulated by alterations in dosage of HP1-Hsalpha through microarray and candidate gene analyses. Understanding the regulation and function of HP1-Hsalpha in breast cancer invasion and metastasis will increase our knowledge of metastatic progression and may lead to new ways to diagnose and treat metastasis.

Presenter: Timothy Moss

Department: Genetics/Biochemistry/MSTP

Category: Graduate Student (College of Medicine)

Graduate Program: Genetics/MSTP

Advisor's Department: Biochemistry

Author 1: Timothy J. Moss

Department: Biochemistry

Other: Genetics/MSTP

Author 2: Sara L. Sloat

Department: Biochemistry

Author 3: Mary J. C. Hendrix

Other: Childrens Memorial Research Center, Northwestern University

Author 4: Naira Margaryan

Other: Childrens Memorial Research Center, Northwestern University

Author 5: Dawn A. Kirschmann

Other: Childrens Memorial Research Center, Northwestern University

Author 6: Lori L. Wallrath

Department: Biochemistry

DOD Era of Hope
June, 2005
Philadelphia, PA

HETEROCHROMATIN PROTEIN 1 AS A CANDIDATE BREAST CANCER METASTASIS SUPPRESSOR

Timothy J. Moss, Laura E. Norwood, Naira Margaryan, Sara L. Cook, Lindsay Wright,
Mary J.C. Hendrix, Dawn A. Kirschmann, And Lori L. Wallrath
Department of Biochemistry, University of Iowa, Iowa City, Iowa
Medical Scientist Training Program, University of Iowa, Iowa City, Iowa
Children's Memorial Hospital, Northwestern University, Chicago, Illinois
lori-wallrath@uiowa.edu

An estimated 200,000 American women are diagnosed with invasive breast cancer each year. The five-year survival rate for women with primary breast cancer is 97%, but the rate is only 23% for women diagnosed with metastatic breast cancer. Currently, the molecular mechanism of metastatic progression is not well understood. We have determined that heterochromatin protein 1, HP1-Hsalpha, is down-regulated in highly invasive/metastatic breast cancer compared to poorly invasive/non-metastatic breast cancer. HP1-Hsalpha is one of three HP1 family members in mammals. HP1 proteins are structural components of centric heterochromatin and also regulate gene expression through epigenetic mechanisms. HP1 family members have two domains, the chromo domain (CD) and the chromo shadow domain (CSD), separated by a flexible hinge region. The CD binds to methylated lysine nine of histone H3, an epigenetic mark for gene silencing. The CSD homodimerizes and generates a platform for interaction with a variety of nuclear proteins containing a PxVxL motif. We hypothesize that HP1-Hsalpha functions as a breast cancer metastasis suppressor by regulating genes involved in metastasis. A highly invasive/metastatic breast cancer cell line, MDA-MB-231, and a poorly invasive/non-metastatic cell line, MCF-7, were used as models to study the role of HP1-Hsalpha in invasion and metastasis. Reduced levels of HP1-Hsalpha in MDA-MB-231 cells is partially due to transcriptional down-regulation by an E-box regulatory element within the promoter region. Introduction of exogenous HP1-Hsalpha into MDA-MB-231 cells reduces in vitro invasion by 29%. An amino acid substitution in the CSD (I165E) that disrupts HP1-Hsalpha homodimerization does not suppress invasion of MDA-MB-231 cells, but an amino acid substitution in the CSD (W174A) that disrupts the platform needed for interaction with proteins containing the PxVxL motif suppresses in vitro invasion by 32%. These results indicate that HP1-Hsalpha dimerization is required for suppression of invasion. Consistent with these findings, introduction of RNAi constructs causing knock-down of HP1-Hsalpha in MCF-7 cells increases their in vitro invasion by 50%. Current studies include the identification of genes misregulated by alterations in dosage of HP1-Hsalpha through microarray and candidate gene analyses. Understanding the regulation and function of HP1-Hsalpha in breast cancer invasion and metastasis will increase our knowledge of metastatic progression and may lead to new ways to diagnose and treat metastasis.